Riegel's Handbook of Industrial Chemistry

NINTH EDITION

VOLUME 1

Edited by James A. Kent

Riegel's

Handbook of Industrial Chemistry

NINTH EDITION

Riegel's

Handbook of Industrial Chemistry

NINTH EDITION

Edited by

James A. Kent, Ph.D.

SPRINGER SCIENCE+BUSINESS MEDIA, LLC

Copyright © 1992 by Springer Science+Business Media New York
Originally published by Van Nostrand Reinhold, New York NY in 1992
Softcover reprint of the hardcover 9th edition 1992

ISBN 978-1-4757-6433-8 ISBN 978-1-4757-6431-4 (eBook)
DOI 10.1007/978-1-4757-6431-4

All rights reserved. No part of this book covered by the copyright hereon may be reproduced or used in any form or by any means--graphic, electronic, or mechanical, including photocopying, recording, taping, or information storage and retrieval systems--without the written permission of the publisher.

3 4 5 6 7 8 9 XXX 01 00 99 98 97 96

Library of Congress Cataloging-in-Publication Data

Riegel, Emil Raymond, 1882--1963.
[Handbook of industrial chemistry]
Riegel's Handbook of industrial chemistry.--9th ed. / edited by James A. Kent.
p. cm.
Includes bibliographical references and index.
1. Chemistry, Technical. I. Kent, james Albert, 1922-- . II. Title. III. Title: Handbook of industrial chemistry.
TP145.R54 1992 92-22660
660--dc20 CIP

To my Wife
ANITA

Contents

Preface		ix
Chapter 1	Economic Aspects of the Chemical Industry	1
Chapter 2	Pollution Prevention and Waste Minimization	15
Chapter 3	Industrial Wastewater and Hazardous Materials Treatment Technology	31
Chapter 4	Applied Statistical Methods and the Chemical Industry	83
Chapter 5	Safety Considerations in the Chemical Process Industries	118
Chapter 6	Managing an Emergency Preparedness Program	176
Chapter 7	Wood and Wood Products	207
Chapter 8	Animal and Vegetable Fats, Oils and Waxes	273
Chapter 9	Sugar and Other Sweeteners	315
Chapter 10	Phosphorus and Phosphates	347
Chapter 11	Fertilizers	367
Chapter 12	Salt, Chlor-Alkali, and Related Heavy Chemicals	408
Chapter 13	Industrial Gases	442
Chapter 14	Sulfur and Sulfuric Acid	458
Chapter 15	Petroleum and Its Products	480
Chapter 16	Natural Gas	510
Chapter 17	Coal Technology	527
Chapter 18	Rubber	598
Chapter 19	Synthetic Resins and Plastics	623
Chapter 20	The Chemistry of Structural Adhesives	708
Chapter 21	Manufactured Textile Fibers	735
Chapter 22	Synthetic Organic Chemicals	800
Chapter 23	Dye Application, Manufacture of Dye Intermediates and Dyes	863
Chapter 24	Industrial Fermentation: Principles, Processes, and Products	916
Chapter 25	The Pharmaceutical Industry	987
Chapter 26	Soap, Fatty Acids, and Synthetic Detergents	1012
Chapter 27	Pigments, Paints, Polymer Coatings, Varnishes, Lacquers, and Printing Inks	1050
Chapter 28	Synthetic Nitrogen Products	1068
Chapter 29	The Agrochemical Industry	1141
Chapter 30	Chemical Explosives and Rocket Propellants	1186
Chapter 31	The Nuclear Industry	1226
Index		1277

Preface

The aim of this book is to present in a single volume an up-to-date account of the chemistry and chemical engineering which underlie the major areas of the chemical process industry. This most recent edition includes several new chapters which comprise important threads in the industry's total fabric. These new chapters cover waste minimization, safety considerations in chemical plant design and operation, emergency response planning, and statistical applications in quality control and experimental planning. Together with the chapters on chemical industry economics and wastewater treatment, they provide a unifying base on which the reader can most effectively apply the information provided in the chapters which describe the various areas of the chemical process industries.

The ninth edition of this established reference work contains the contributions of some fifty experts from industry, government, and academe. I have been humbled by the breadth and depth of their knowledge and expertise and by the willingness and enthusiasm with which they shared their knowledge and insights. They have, without exception, been unstinting in their efforts to make their respective chapters as complete and informative as possible within the space available. Errors of omission, duplication, and shortcomings in organization are mine.

Grateful acknowledgment is made to the editors of technical journals and publishing houses for permission to reproduce illustrations and other materials and to the many industrial concerns which contributed drawings and photographs.

Comments and criticisms by readers will be welcome.

James A. Kent
Stevensville, Maryland

1

Economic Aspects of the Chemical Industry

F. E. Bailey, Jr.* and J. V. Koleske**

Within the formal departments of science at the traditional university, chemistry has grown to have a unique status because of its close correspondence with an industry and a branch of engineering—the chemical industry and chemical engineering. There is no biology industry, but drugs, pharmaceuticals, and agriculture are closely related disciplines. There is no physics industry although power generation, electricity, and electronics industries do exist. But connected with chemistry, there is an industry. This unusual correspondence probably came about because in chemistry one makes things from basic raw materials—chemicals—and the science and the use of chemicals more or less grew up together during the past century.

Since there is a chemical industry, which serves a major part of all industrialized economies, providing in the end synthetic drugs, fertilizers, clothing, building materials, paints, elastomers, etc., there is also the subject of "chemical economics"; and it is this subject, the economics of the chemical industry, that is the concern of this chapter.

DEFINITION OF THE CHEMICAL INDUSTRY

Early in the twentieth century, the chemical industry was considered to have two parts: the manufacture of inorganic chemicals and the manufacture of organic chemicals. Today, the Standard Industrial Classification (SIC Index) of the United States Bureau of the Census defines "Chemical and Allied Products" as comprising three general classes of products: "(1) basic chemicals such as acids, alkalis, salts, and organic chemicals; (2) chemicals to be used in further manufacture such as synthetic fibers, plastics materials, dry colors, and pigments; and (3) finished chemical products to be used for ultimate consumer consumption as architectural paints, drugs, cosmetics, and soaps or to be used as materials or supplies in other industries such as industrial paints, adhesives, fertilizers, and explosives." [1] An even broader description that is often considered is that of the "chemical

*Deceased.
**Consultant, 1513 Brentwood Rd., Charleston, WV 25314.

process industries," major segments of which include: chemical and allied products and petrochemicals; pulp and paper; petroleum refining; rubber and plastics; and stone, clay, and glass products.

THE PLACE OF THE CHEMICAL INDUSTRY IN THE ECONOMY

Because the chemical industry is a major sector of any advanced national economy, a forecast of trends in the chemical industry must fall within certain general guidelines that are established by the national economy. A forecast for the chemical industry in the United States must be within the general boundaries set for the overall social and economic forecasts for the country.

It has been clear for many years that certain demographic and societal issues would have a dominant effect on the U.S. economy of the 1990s. It was evident, for example, that from the late 1980s through the year 2000 there would be a decline in the growth of the work force in the United States. There is a direct relation between the growth of the work force and the growth of the GNP (Gross National Product, the sum of all goods and services produced in a year); the decline in the growth of the work force in this period was determined by the number of women in the usual child-bearing age group (18 to 35) and by family-size decisions made in the 1960s. Therefore, this factor was set and calculable from census data obtained 20 years earlier.

A predicted decline in the growth of the GNP due to a declining work force can be offset if worker productivity increases or if the age of retirement from the work force is extended beyond the usual retirement age of 65. In the United States during the 1980s, there was little overall increase in worker productivity; and, in many industries, there was a move toward retirement before age 65. In many industries, the average age of retirement moved toward 60 from the long-accepted norm of 65. These trends, as well as a marked growth in the service sector of the economy as opposed to such growth in the manufacturing sector, have contributed to a shortage of technically trained chemists and engineers entering the work force. This shortage is expected to become severe in the latter part of the 1990s.

Another factor in the overall economy that must be considered is that the GNP represents a "volume" of goods and services measured in value or "dollars." This "dollar volume" is very sensitive to inflation (or deflation). To remove this sensitivity for forecasting purposes, GNP is expressed in "constant dollars," dollars "deflated" by the annual inflation rate to some base year such as 1982.

Against this brief discussion of the general demographic, societal, and economic factors that govern forecasting, Table 1.1 gives a general picture of the economy of the United States in terms of the GNP and chemical industry production.

The forecast for the early part of the 1990s, which is subject to all the uncertainties and unpredictabilities of social and international events, is for steady growth of the U.S. chemical industry, but growth at a rate lower than that experienced in the mid-1980s and moderated by an average annual inflation rate of about 5 percent per year. This forecast includes a gradual increase in the price of imported oil and petroleum products during the early 1990s compared with lower imported oil costs in the late 1980s. Changes such as that caused by the 1990–91 Persian Gulf crisis have not been considered although such events certainly can have a profound effect on prices. At present the price of oil appears to be fairly stable, at least for the short term.

This forecast is presented to highlight the sensitivity and the place of the chemical industry in the national economy. The importance of the cost of imported raw materials such as petroleum products in both the economy and the forecast highlights two important concerns. One is that the chemical industry is worldwide and interconnected; not only does the United States both import and export a wide variety of raw materials and chemical products, but major United States–based chemical companies have manufacturing and sales facilities abroad, and a large number

TABLE 1.1 U.S. Economy and the Chemical Industry[2]

Year	United States GNP (Current dollars, billions)	United States GNP (1982 dollars, billions)	Percent Change in Chemicals and Allied Products Annual Production Index
1987	4500	3800	6.1
1988	4900	4000	8.5
1989	5200	4100	4.9
1990 (estimate)	5500	4200	2.4
1995 (forecast)	8000	4700	2.8*

*Annual average for 1990–95.

of foreign-based companies have manufacturing and sales facilities in the United States. The other is that the United States economy is dependent on the balance of trade, the difference between the dollar value of exports and that of imports. A negative trade balance means that dollars spent abroad to import goods and services exceed the value of goods and services exported, a circumstance that effectively increases the cost of goods and services purchased in the United States—a net inflationary effect. To a large extent during the 1980s, this potentially inflationary effect was offset by foreign investment in the United States; however, this offsetting of a negative trade balance by investment cannot be depended on to continue through the 1990s. At some point, foreign investors will demand a return on their investment, in effect an export of dollars, with an inflationary result on the domestic economy.

In foreign trade, the performance of the chemical industry of the United States has been outstanding. While the overall balance of trade has been negative, the chemical industry has been one of the truly strong sectors in the U.S. economy; see Table 1.2. If the total world export market for chemicals is considered, that is, the sum of all of the chemicals exported by all the world's national economies, the U.S. chemical industry held about a 15 percent market share during the decade of 1979 to 1989.

Major segments in the U.S. chemical trade balance in 1989 were:

Organic chemicals— + $3.6 billion
Plastics materials— + $4.6 billion
Pharmaceuticals— + $1.6 billion

The less favorable overall trade balance of the United States was due principally to imports of manufactured goods and petroleum products; see Table 1.3.

To support the U.S. chemical economy in 1989, there was a work force of more than one million, 600,000 of whom were occupied in chemical production. The largest single sector of this work force, about 20 percent,

TABLE 1.2 U.S. Balance of Trade

	Total Trade Balance (billions of dollars)			Chemical Trade (billions of dollars)		
Year	Export	Import	Balance	Export	Import	Balance
1979	181	206	−25	17	7	+10
1987	253	406	−153	26	16	+10
1989	364	473	−109	36	20	+16
1990	394	495	−101	39	22	+17

TABLE 1.3 U.S. Trade Balance in 1989 (billions of dollars)

Sector	*Import*	*Export*	*Balance*
Chemicals	20	36	+16
Machinery	206	148	−58
Other manufactured goods	139	60	−79
Fuels and petroleum products	53	9	−44

was involved in the manufacture of ethical drugs and pharmaceuticals. This group was closely followed by those workers involved in the manufacture of plastics and synthetic materials. To back up this immensely important part of the economy, the chemical industry (chemicals and allied products) employed about 75,000 scientists and engineers, which is about 10 percent of the total number of scientist and engineers employed by industry in the United States. On an average, the larger chemical companies (top 15 companies in chemical sales) invested approximately 4 percent of their gross sales in research and development activities to support growth and to maintain their market share.

CHARACTERISTICS OF THE CHEMICAL INDUSTRY

Investment Trends

The chemical industry tends to be a high-investment business. Capital spending by the chemical and allied products industry in the United States has been a sizable percentage of the total spent for all manufacturing. Amounts spent in this industry and in certain facets of it are detailed in Table 1.4. The amount spent for all chemical process industries has been, of course, even larger; it totaled $68.9 billion in 1990, for example, a 5.3 percent increase over the amount spent in 1989. For perspective, annual expenditures for new plant and equipment in the United States for the chemical and allied products industry in recent years have averaged about 2.5 times

TABLE 1.4 New Capital Spending in the U.S. Chemical and Allied Products Industry and Comparison with That Spent a Decade Earlier[1]

	Billions of Dollars	
	1986	*1976*
Total chemical and allied products	7.9	7.1
Selected Segments		
Agricultural chemicals	0.52	1.05
Industrial inorganic chemicals	0.81	0.81
Industrial organic chemicals	1.84	2.69
Plastic materials, synthetic resins, and thermoplastic elastomers	1.90	1.37
Soap, detergents, perfumes, cosmetics, and other toilet products	0.74	0.29
Miscellaneous chemicals	0.51	0.32

the amount spent for iron and steel and about half of that invested in the petroleum industry. For the past decade a significant part of these capital investments have been made in pollution control and projects related to the environment.

Much of the capital investment in the chemical industry is spent for facilities used to produce major chemicals (Table 1.4) in truly enormous quantities. The volume produced is reflected in the size of plants being built to achieve the required economies of scale. That such economies are achieved is seen in the more modest increases in the chemical producers' price indices relative to the inflation levels in the general economy. (Economy of scale refers to the relative cost of building a larger plant; a rule of thumb is that the relative cost of building a smaller or a larger plant is the ratio of the productivities of the two plants being considered, raised to the 0.6 power. In other words, the unit cost of producing a chemical markedly decreases as the size of the plant producing it is increased, provided that the plant can be operated near capacity.)

Today, a typical, base petrochemicals plant will consume the equivalent of 30,000 barrels per day of naphtha to produce about one billion pounds of ethylene a year, plus 2.5

TABLE 1.5 Employment in Selected Parts of the Chemical Industry in 1990, 1987, and 1977[2]

	Thousands of Employees		
Chemical Industry	*1977*	*1987*	*1990*
Chemical and allied products	1,074	1,026	1,086
Segments			
Cosmetics, perfumes, etc.	53	70	N/A*
Drugs	181	212	239
Industrial inorganic chemicals	162	135	135
Industrial organic chemicals	166	152	153
Paints	66	63	63
Pharmaceutical preparations	145	169	N/A
Plastic materials	83	76	N/A

*N/A = not available.

billion pounds of coproducts. To be economically feasible, for example, plants for the production of monomers such as vinyl chloride and styrene for plastic products must be scaled in the billion-pound-per-year range.

Along with these very large plants and the associated enormous investment in them, most of the chemical industry is characterized by high investment versus low labor components in the cost of manufacture. The National Industrial Conference Board statistics list the chemical industry as one of the highest in terms of capital investment per production worker. The investment per worker in a base petrochemicals olefins plant may well exceed a quarter of a million dollars. Once again, however, such an index covers a spectrum of operations, and for a profitable chemical specialties manufacturer the investment may be on the order of 25,000 dollars per worker. Employment in selected parts of the chemical industry is given in Table 1.5.

Commercial Development and Competition Factors

In an earlier period of the chemical industry's development, chemical companies were generally production-oriented, exploiting a process to produce a chemical and then selling it in rapidly expanding markets. The plant sizes and investments required for participation were small fractions of what is needed to participate today. Raw materials often were purcharsed to produce chemical intermediates for sale. Small plants operating in small manufacturing complexes did not present the obvious problems of environmental pollution, which everyone has become increasingly aware of during the past decade. A new investment in chemical production today must include a significant proportion of the total outlay for abatement and control of environmental intrusion.

As the industry has grown, there has been a strong tendency toward integration, both forward and back. Petroleum producers have found opportunities based on their raw materials position to move into chemical manufacturing. Chemical companies, on the other hand, have moved to assure their access to low-cost raw materials. Similarly, producers of plastic materials have moved forward to produce fabricated products, such as films, fibers, and consumer items, while fabricators have installed equipment to handle and formulate the plastic materials to provide a supply at the lowest possible cost.

With ever higher investment needs and increasing cross-industry competition, much greater sophistication has been required in marketing analysis and selection of investment opportunity. The enormity of the investment now required for successful participation does not allow multiple approaches for the private investor. Consequently, a high degree of market orientation tends to predominate in the chemical industry, along with increasingly focused research and development programs.

A major trend in industrial chemistry has been an emphasis on improved processes for the production of major chemicals. The need for higher-efficiency, lower-cost processes has been accentuated by relatively slow growth rates in overall production of major industrial chemicals in recent years, as described in Table 1.6 for inorganic chemicals and Table 1.7 for organic chemicals. Thus profitability growth has had to be achieved from higher efficiencies in production.

For plastics materials, there also was relatively slow growth—except for a few bright spots such as high-density polyethylene

TABLE 1.6 U.S. Production of Major Inorganic Chemicals in 1990 and the Average Annual Production Growth Rate, 1980–90[3]

Chemical	*1990 Production Amount, 1980–90 (billion of pounds)*	*Annual Growth Rate (percent per year)*
Sulfuric acid	88.6	0
Nitrogen	57.3	5.1
Oxygen	39.0	0.9
Lime	34.8	−0.9
Ammonia	33.9	−1.5
Phosphoric acid	24.4	1.2
Sodium hydroxide	23.4	0.1
Chlorine	21.9	−0.4
Sodium carbonate	19.9	1.8
Urea	15.8	0.1
Nitric acid	15.5	−1.7
Ammonium nitrate	14.2	−2.5
Carbon dioxide	11.0	6.2
Ammonium sulfate	5.0	1.6
Hydrochloric acid	4.7	−2.1
Potash (K_2O basis)	3.6	−3.1
Carbon black	2.9	1.2
Aluminum sulfate	2.4	−0.6
Titanium dioxide	2.2	4.2
Sodium silicate	1.8	0.9
Sodium sulfate	1.5	−4.3
Calcium chloride	1.4	−3.5

TABLE 1.7 U.S. Production of Major Organic Chemicals in 1990 and Average Annual Growth Rate, 1980–90[3]

Chemical	*1990 Production Amount, 1980–90, (billion of pounds)*	*Annual Growth Rate (percent per year)*
Ethylene	37.5	2.7
Propylene	22.1	4.9
Ethylene dichloride	13.3	1.8
Benzene	11.9	−2.2
Vinyl chloride	10.7	5.1
Ethyl benzene	9.0	1.6
Styrene	8.0	1.6
Methanol	8.0	1.1
Terephthalates	7.7	2.4
Formaldehyde	6.4	1.4
Toluene	6.1	−1.9
Xylene	5.7	−1.4
Ethylene oxide	5.6	0.7
Ethylene glycol	5.0	1.4
Phenol	3.5	3.2
Acetic acid	3.5	2.4
Propylene oxide	3.2	6.1
Butadiene	3.2	1.2
Acetone	2.2	0.7
Vinyl acetate	2.6	2.8
Cyclohexane	2.5	2.3

TABLE 1.8 U.S. Production of Major Plastics Materials in 1990 and Annual Growth Rate, 1980–90[3]

Plastic Material	*Annual Production (billion of pounds)* 1990	1980	*Production Growth Rate* 1980–90
Polyethylene			
Low density	11.2	7.3	1.9
High Density	8.3	4.4	6.6
Poly(vinyl chloride)*	9.1	5.5	5.2
Polypropylene	8.3	3.7	8.6
Polystyrene	5.0	3.5	3.6
Unsaturated polyester	1.2	1.0	2.6
Epoxide	0.50	0.32	5.0

*Includes copolymers of vinyl chloride.

and polypropylene—over the decade of 1979–89; see Table 1.8. Toward the end of the 1980s, growth decreased as increasing emphasis was placed on the recycling of packaging materials, and as new construction, a major use area for plastics materials, declined. As the 1990s began, recycling was showing strong growth, and there was no reason to think recycling would decrease during the nineties.

In the plastic and synthetic materials sector, in which useful fabricated articles are produced directly, for example, from polymers such as polyethylene, polypropylene, and poly(vinyl chloride), there is now major emphasis on the reuse of plastic instead of discarding it, particularly as packaging of various sorts, in trash for landfill. The major problems are not so much technical as they are ones of logistics and the development of a new kind of infrastructure to accommodate recycling.

Some 20 years ago, there was a proposal in Germany to reuse "virgin" poly(vinyl chloride) blow-molded beer bottles sold in the larger football (soccer) stadiums by regrinding the plastic bottles and extruding the ground product as pellets for use in flooring applications. The concept was that a large number of empty bottles could be collected at one point and reprocessed essentially "on the spot." This proposal was not successful; but it illustrates the major recycling problems of collection, sorting by material, and reprocessing for reuse. The proposal also presaged many active new ventures being started in the 1990s, such as the Du Pont effort to reprocess polyethylene and polyester polymers into fencing, fence posts, and garden furniture, and the new ventures by Mobil Chemical and Union Carbide to reprocess polyethylene packaging, film, and wrap. American National Can Company has demonstrated, with the help of the Center for Plastics Recycling Research at Rutgers University, that a blow-molded, layered polypropylene bottle could not only be reprocessed into shampoo and detergent bottles but also into automobile bumper fascia. Today supermarkets are collection depots for plastic grocery bags, and many municipalities are requiring separation of trash into recyclable components. Even more desirable are the voluntary separation and proper disposal of selected trash items that have been undertaken by many individuals.

The principal recycling problems that must be overcome to make plastic-material recycling feasible on a significant scale include the development of commercial pathways for consumer disposal, collection, and reprocessing. A technical problem remains in the collection–reprocessing step of sorting by material type. In the examples of the poly(vinyl chloride) bottle for stadium beer in Germany and polyethylene shopping bags and wrap, identification and collection are relatively simple if the collection is done at the point of sale, or if material is collected later at the place of original sale. However, it must be realized

that different polymeric materials usually are not compatible with one another. Two incompatible materials generally will not mix or blend together sufficiently to yield a product with any utilizable strength properties, and are useful only for their fuel value. The process of sorting items into separate materials (e.g., poly(vinyl chloride), polyester, polypropylene), by the consumer–home owner or at a community-operated collection center, is still difficult. A recent solution to this problem has been to mark shaped plastic articles with a code that permits easy identification by classification or type.

Technological Orientation

The chemical industry is a high-technology industry, even though it is now more marketing-oriented and competitive than it was in its earlier period of development. This orientation is shown by the number of scientists and engineers employed in research and development in the chemical industry relative to other industries (Table 1.9). In general, the chemical industry is among the largest employers of scientists and engineers, and it puts a sizable percentage of the total U.S. business investment in research and development.

The contemporary scientist or engineer engaged in research and development in the chemical industry represents individually a high-investment occupation. Since the mid-1950s, chemistry has become increasingly an instrumental science. The instruments now routinely used are both highly sophisticated and costly. A major research project would not be undertaken today without access to a variety of spectrophotometers, spectrometers, chromatographs, etc., as well as the necessary physical/chemical instruments for molecular structure determinations and reaction kinetics. Pilot plants are highly automated and instrumented. Both the basic researcher and the pilot plant engineer require access to computer facilities. In 1978, the average annual cost to maintain an operating R&D scientist or engineer in the chemical industry was about $75,000. Today, it may be three or four times that amount. Impressive as these statistics may be in representing the business investment in chemicals R&D in the United States, R&D spending in the chemicals industry as a percent of sales declined from about 4 percent in 1970 to about 2.5 percent in 1980, although today it remains about 4 percent in the largest chemical companies. This relatively low level is a concern because reinvestment in R&D in other industrialized countries, particularly in West Germany and Japan, has remained at a higher level as measured by this index.

TABLE 1.9 Scientists and Engineers in Research and Development in the United States in 1990 and a Decade Earlier[2]

	Thousands of Employees		
	1990	*1987*	*1980*
Chemicals and Allied Products	78.4	75.2	51.4
Segments			
Industrial Chemicals	23.0	22.4	20.9
Drugs	33.0	32.6	21.6
Other Chemicals	22.4	20.2	8.9

Obsolescence and Dependence on Research

The high technology level that characterizes the chemical industry, and which is reflected in heavy investments in R&D, generally concerns discovery and development of new products and improvements in the manufacture of known products. The first area is more conspicuous: the pharmaceutical for a specific disease; the narrow-spectrum, transient pesticide; the new superperformance, composite system for an internal combustion engine; or a thermoset polymer/graphite fiber-composite material for a high-technology military aircraft, such as the "stealth" bomber or fighter. The second area, however, makes viable the circumstances outlined earlier, where increasing investments can be made to produce larger quantities of materials. The development of a new, lower-cost process for a commercial product can permit the development of a profitable opportunity, or can spell

disaster for a company with existing investment in a now-obsolete plant. Major reductions in manufacturing cost can be achieved, for example, by reducing the number of reaction steps required, changing to a lower-cost or more available raw material, or eliminating coproducts, costly separations, and environmental intrusions. The ability of a process scheme to contain or avoid a pollutant can be a deciding factor in the continuance of a manufacturing operation. Examples of the above situations will make the economic consequences clear.

Acetic acid production in the United States has increased about eightfold in the last 40 years. From the 1930s, acetic acid was produced by a three-step synthesis from ethylene; acid hydrolysis to ethanol, then catalytic dehydrogenation to acetaldehyde, then direct liquid-phase oxidation to acetic acid and acetic anhydride as coproducts:

$$CH_2{=}CH_2 \xrightarrow{H_2SO_4/H_2O} C_2H_5OH$$

$$\xrightarrow{Cu/Cr} CH_3CHO$$

$$CH_3CHO \xrightarrow{[O],Co}$$

$$CH_3CO{-}O{-}OCCH_3 + CH_3COOH$$

In the 1940s a major process change was introduced—direct oxidation of butane to acetic acid and coproducts (such as methylethylketone):

$$C_4H_{10} \xrightarrow{[O]} CH_3COOH$$

$$+ CH_3COC_2H_5 + \text{others}$$

Having fewer steps in the synthesis was reflected in lower cost and investment. In 1969, another advance was announced, synthesis of acetic acid from methanol and carbon monoxide with essentially no coproducts:[4,5]

$$CH_3OH + CO \xrightarrow{I/Rh} CH_3COOH$$

The absence of coproducts reduces production costs and investment in distillation and other separation systems. These are very attractive process features in an industry where the principally accepted measure of business quality is return-on-investment.

Acetic anhydride is required as a process intermediate in acetylations. To obtain acetic anhydride from acetic acid, acetic acid is first pyrolyzed to ketene, which then reacts with recovered acetic acid to yield the anhydride:

$$CH_3COOH \xrightarrow{Heat} CH_2{=}C{=}O$$

$$CH_2{=}C{=}O + CH_3COOH \longrightarrow$$

$$CH_3CO{-}O{-}OCCH_3$$

In 1980, the Tennessee Eastman unit of Eastman Kodak announced that it would begin construction of a plant to make acetic anhydride from coal.[6,7] This decision reflected a changing of the raw materials base of much of the chemical industry due to such factors as the rising cost of natural gas and petroleum and the large coal reserves of the United States.

In the new process, synthesis gas (carbon monoxide and hydrogen) is made from coal. Then, methanol is produced from the synthesis gas. (Previously, methanol had been produced chiefly from natural gas methane.)

$$CO + 2H_2 \rightarrow CH_3OH$$

Methanol can then react with acetic acid to give methyl acetate:

Acetic anhydride is then obtained from the catalytic carbonylation of methyl acetate with carbon monoxide:[5]

$$CH_3COO{-}CH_3 + CO \rightarrow$$

$$CH_3CO{-}O{-}OCCH_3$$

The attractiveness of this process is twofold: (1) the raw materials base of synthesis gas from coal and (2) the avoidance of energy-consuming manufacture of ketene by pyrolyzing acetic acid.

The increase in the production of vinyl chloride, the principal monomer for poly(vinyl chloride) plastics, which are used in vinyl flooring, phonograph records, shower curtains, raincoats, carseat upholstery, house siding, pipe, and so on, has been even more spectacular. Production in the United States has increased from 250 million pounds in 1950 (when it was declared by many industry economic forecasters to be a mature chemical commodity) to over one billion pounds in 1960, to about 3.5 billion pounds in 1970, to over 7 billion pounds in 1980, and to almost 11 billion pounds in 1990.

During the early development period of vinyl polymers in the 1930s, vinyl chloride was produced via catalytic addition of hydrogen chloride to acetylene:

$$CH{\equiv}CH + HCl \xrightarrow{HgCl} CH_2{=}CHCl$$

Later, a so-called balanced process was introduced, in which, by addition of chlorine to ethylene, ethylene dichloride was produced:

$$CH_2{=}CH_2 + Cl_2 \xrightarrow{[O]/Cu} CH_2ClCH_2Cl$$

Ethylene dichloride could then be cracked to vinyl chloride and HCl, with the hydrogen chloride recycled to produce vinyl chloride from acetylene:

$$CH_2ClCH_2Cl \xrightarrow{Heat} CH_2{=}CHCl + HCl$$

At that time, vinyl chloride was being produced from chlorine, acetylene, and ethylene. More recently, catalytic oxychlorination has been developed, in which vinyl chloride is produced from ethylene and hydrogen chloride:[8]

$$CH_2{=}CH_2 + HCl \xrightarrow{[O]/Cu} CH_2{=}CHCl + H_2O$$

The hydrogen chloride can be obtained via cracking of ethylene dichloride. The oxychlorination process freed vinyl chloride from the economics of the more costly raw material, acetylene. (Deliberate acetylene manufacture is energy-intensive. Although by-product acetylene from gas cracking is less expensive than the raw material, it has not been available in sufficient supply for the large, near billion-pound-per-year, vinyl chloride units.)

During the long period of development of poly(vinyl chloride) into one of the major plastics material, several basic processes for making PVC evolved. In all of these processes vinyl chloride was handled as a liquid under pressure. Despite the relative ease with which it could be polymerized by free radical initiators, the monomer, vinyl chloride, was regarded as an innocuous, relatively inert chemical. A number of producers of PVC resins were caught by total surprise in the 1970s when it was found that long-term (20-year) exposure to vinyl chloride monomer could cause rare forms of tumors.[9]

During the 1960s, vinyl chloride sold in the United States for five to six cents per pound. In the presence of traces of air (oxygen), it would form low concentrations of peroxide, which could collect in compressors and on occasion rapidly decompose to blow out compressor seals. Rather than recover and compress the inexpensive monomer for recycle from stripping and drying operations at the end of the polymerization process, some manufacturers vented it into the atmosphere. After the discovery that vinyl chloride was a carcinogen, venting was not permissible; containment and recovery were mandatory. Some older processes and manufacturing facilities could not be economically modified to incorporate monomer containment, so operations were discontinued. This case is but one example of the impact that necessary environmental controls can have on manufacturing processes and operations.

Propylene oxide is another basic chemical, used in manufacturing intermediates for urethane foams (used in cushioning and insulation) and brake and hydraulic fluids. The volume of propylene oxide produced increased from 310 million pounds in 1960 to 3.2 billion pounds in 1990. The classical industrial synthesis has been the reaction of chlorine with propylene to produce the chlorohydrin, followed by dehydrochlorination with caustic to produce the epoxide, propylene oxide, plus salt. In this case, both

the chlorine and the caustic used to effect this synthesis are discarded as a valueless salt by-product:

$$CH_3CH{=}CH_2 + Cl_2 + H_2O \longrightarrow CH_3CH(OH)CH_2Cl$$

$$CH_3CH(OH)CH_2Cl \xrightarrow{\text{Caustic}} CH_3CH\underset{O}{\diagup\!\!\diagdown}CH_2 + \text{salt}$$

A more economical process has been commercialized. In one version, a hydroperoxide is produced by catalytic air-oxidation of a hydrocarbon such as ethylbenzene. Reaction of this hydroperoxide with propylene yields propylene oxide as a coproduct. This direct peroxidation can be carried out with other agents to give different coproducts such as *t*-butanol or benzoic acid.[10,11]

$$\text{Aryl-}C_2H_5 \xrightarrow{[O]/V} \text{Aryl-}CH(CH_3)OOH$$

$$\text{Aryl}{-}CH(CH_3)OOH + CH_3CH{=}CH_2 \longrightarrow CH_3\text{-}CH\underset{O}{\diagup\!\!\diagdown}CH_2 + \text{Aryl}{-}CH(CH_3)OH$$

When the economics are balanced, a significant cost reduction is achieved by eliminating the coproduct salt, which is of low value and presents a disposal problem. Further, a process can be designed to produce a coproduct that can be used or sold as a chemical intermediate. In the case of using isobutane as the starting hydrocarbon, the by-product is *t*-butanol, which can then be converted to methyl *t*-butyl ether, which is the gasoline additive used to replace lead in "lead-free" gasoline.

If a company is in the business of making and selling products such as acetic acid, vinyl chloride, propylene oxide, or other chemicals and has plans to stay in business and to expand its facilities to serve growing markets, it at least must have economically competitive processes. Today this means being competitive with not only any new processes developed in the United States, but also with any new process technology developed in Western Europe, Japan, and Russia—for the chemical industry is a worldwide industry. This is readily apparent from the data in Tables 1.10, 1.11, and 1.12, which describe the sales for the largest chemical producers in the United States, Western Europe, and Japan, respectively. Further, the processes that are operative must be environmentally compatible—all toxic or carcinogenic by-products or waste must be contained and disposed of harmlessly. Even a relatively innocuous by-product such as salt must be disposed of so as not to intrude on the environment.

The profound effect of environmental concerns on the manufacture of a chemical is

TABLE 1.10 Largest U.S. Chemical Producers[2,12]

	Chemical Sales Millions of Dollars			
Company	*1969*	*1979*	*1989*	*1990*
Du Pont	*3,655*	*12,572*	*15,249*	*15,571*
Dow Chemical	*1,876*	*9,255*	*14,179*	*14,690*
Exxon	*1,004*	*5,807*	*10,559*	*11,153*
Union Carbide	*2,933*	*9,177*	*7,962*	*7,621*
Monsanto	*1,939*	*6,193*	*5,782*	*5,711*
Hoechst Celanese	*1,250*	*3,010*	*5,658*	*5,499*
General Electric	—	—	*4,929*	*5,167*
Occidental Petroleum	—	—	*5,203*	*5,040*
BASF (U.S.A. only)	—	—	*4,461*	*4,366*
Amoco	—	—	*4,274*	*4,087*

TABLE 1.11 Largest Western European Chemical Producers

Company	Country	Chemical Sales Millions of Dollars 1969	1979	1988
BASF	Germany	2,430	15,018	24,925
Hoechst	Germany	2,550	15,870	23,275
Bayer	Germany	2,550	15,079	22,993
ICI	United Kingdom	3,250	11,389	20,817
Ciba-Geigy	Switzerland	—	—	12,087
Rhone-Poulenc	France	1,840	7,940	10,962
Montedison	Italy	2,620	8,224	10,846
Norsk Hydro	Norway	—	—	9,207
Akzo	The Netherlands	—	—	8,374
Degussa	Germany	—	—	7,730

1988 data from "Facts and Figures for the Chemical Industry," *Chemical and Engineering News*, p. 81, June 19, 1989.

TABLE 1.12 Largest Japanese Chemical Producers

Company	Chemical Sales Millions of Dollars 1979	1988
Asahi Chemical Industry	2,359	6,384
Mitsubishi Kasei	2,967	6,132
Takeda Chemical Industry	—	4,567
Sumitomo Chemical	2,716	4,341
Toray Industries	2,094	4,317
Sekisui Chemical	1,300	3,936
Showa Denko	1,706	3,679
Dainippon Ink & Chemicals	1,416	3,440
Mitsui Toatsu Chemicals	1,767	3,082
Mitsubishi Petrochemical	1,558	2,654

1988 data from "Facts and Figures for the Chemical Industry," *Chemical and Engineering News*, p. 83, June 19, 1989.

TABLE 1.13 Aerosol, Pressurized Product Sales in the United States in 1978

Use	Percent of Sales
Personal products and toiletries	33%
Household products	28
Coatings and finishes	14
Automotive	7
Food products	6
Insect sprays	6
Industrialized products	5
Animal products	1
Miscellaneous	0.5

reflected in the history of aerosol pressurized products. These products are familiar to the consumer in aerosol spray cans containing hair sprays, deodorants, dessert toppings, or insect sprays (Table 1.13). The market for these products grew enormously in the 1960s, with rapid consumer acceptance of packaged convenience products.[13,14]

The aerosol spray product is a pressurized formulation with a propellant gas. During the rapid growth of these products, the major propellant gases were chlorofluorocarbons. Then, in 1973, the uncontrolled release of chlorofluorocarbons into the atmosphere was linked to possible depletion of the ozone layer in the earth's atmosphere. Since stratospheric ozone provides significant protection at the earth's surface from ultraviolet radiation from the sun, depletion of the ozone layer could be forecast to lead to skin cancers, reductions in seafood and grain crops, and alteration of the carbon dioxide level in the atmosphere. Carbon dioxide is thought to be a potential culprit in "global warming."

Popular concern over "global warming" has now spread from its origin mainly in the large developed economies, which were the principal users of halocarbon propellants, to virtually all countries. The release into the atmosphere of contaminants that can affect the ozone layer, particularly the "ozone hole"

over the Antarctic region, has been identified with the very general use of fire-extinguishing devices, refrigerants, and fabricating processes such as the manufacture of foamed plastic. An international protocol formulated in Montreal in 1988 has now been generally accepted, and plans to phase out the manufacture of ozone-depleting chemicals during the next few years are now in place—an extraordinary example of international agreement and cooperation.[15]

THE NEXT VISTA IN CHEMISTRY

While the major thrust in process development for chemicals in the 1980s was to reduce by-product manufacture, to increase process efficiencies, and by these means to control and reduce the impact that chemical processes and products may have on the environment, the next vista may well involve radically new chemical reaction concepts. Enantiomorphic selectivity, which is the ability to control the stereoisomer produced from a complex organic reaction, will be particularly important in the areas of pharmaceuticals and agricultural chemicals, where it is commonly the case that only one of many possible stereoisomers has the desired bioactivity. In 1992, this area of chemical synthesis is still in its true infancy.

A simple example of enantiomorphic selective synthesis can be taken from polymer chemistry and the polymerization of propylene oxide. Propylene oxide contains an asymmetric carbon atom, C*:

The polymerization of propylene oxide can be induced in a number of ways, two of which can be given in this example.[16] With a caustic initiator such as potassium hydroxide, polymerization does not discriminate between the stereoisomers, and a random distribution of stereoisomers occurs. The resulting polymer is a liquid or a soft, amorphous solid, depending on molecular weight, that softens at about −40°C. This material can be the main structural unit in specialty polyether elastomers and in many flexible urethane foams found in seat cushions, for example.

If, however, a complex initiator based on aluminum or zinc alkyls is used, a product with identical chemical constitution but entirely different physical characteristics results. The reason for this is that the individual polymer chains are produced from either one or the other of the two stereoisomers. The polymer formed from the racemic mixture of *d*- and *l*-propylene oxides is a solid thermoplastic with a crystalline melting point of about +75°C. An even more dramatic result was obtained by Vandenberg[17] with the polymerization of 2,3-epoxybutane:

in which there are two asymmetric carbon atoms, and diestereoisomer selectivity is possible. Vandenberg was able to produce three entirely different polymers from this one monomer, depending on the initiator used:

1. A soft, amorphous rubbery polymer.
2. A crystalline polymer melting at 162°C.
3. A crystalline polymer melting at 100°C.

Examples of this kind of enantiomorphic or chiral selectivity are now being found in organic synthesis. Asymmetric synthesis, for example, has been demonstrated with stereocontrolled Michael addition in the synthesis of beta-lactams using chiral catalysts, where an acyl ligand such as acetyl is bound to cyclopentadiene carbonyl triphenylphosphine.[18] Essentially complete enantiomorphic selectivity has been achieved in this Michael addition synthesis. Another case is enantiomorhic ketone reduction in ethylbenzene reduction in the ethylation of benzaldehyde. Using chiral catalysts, 97% selectivity has been achieved.[19] Closely related research involves the making of catalytic antibodies and hybrid enzymes.[20]

Composites were only briefly mentioned earlier, but this area, which will supply shaped articles to the aircraft/aerospace, industrial, and recreation markets, is expected to grow at a compounded rate of 8 percent from 1988 through the year 2000.[21] In 1988, worldwide consumption of such polymer composites was $2,300 million.

Stereolithography,[22] a new technique for prototyping products, is currently used for modeling and is envisioned as a production tool that will use large quantities of highly upgraded, formulated chemicals. Yet, a few years ago this highly useful technique that couples computers, lasers, or other curing methods and chemistry was unknown, with the first patent being issued in 1986.[23]

It is in areas such as these that the commercial thrusts of the 1990s will have to be aimed if they are to maintain the chemical industry's place in the viable economies of the world.

REFERENCES

1. *Chemical Economics Handbook*, Stanford Research Institute, Menlo Park, CA.
2. "Facts and Figures," *Chemical and Engineering News* (June 17, 1991).
3. "Top 50 Chemical Production Resumed Growth Last Year," *Chemical and Engineering News*, p. 13 (Apr. 8, 1991).
4. Belgian Patent 713,296 (to Monsanto Co.).
5. Parshall, G. W., "Organometallic Chemistry in Homogeneous Catalysis," *Science*, **208**, 1221 (1980).
6. *Chemical Week*, **126** (3), 40 (1980).
7. *Chemical and Engineering News*, **58** (2), 6 (1980).
8. British Patent 1,027,277 (to PPG Industries) and British Patent 1,016,094 (to Toyo Soda).
9. *Chemical Week*, **125** (6), 22 (1979).
10. French Patent 1,460,575 (to Halcon).
11. *Chemical Week*, **127** (14), 47 (1980).
12. *Chemical and Engineering News*, **58** (1), 6 (1980).
13. "Top 100 Chemical Producers Post Another Record Sales Year," *Chemical and Engineering News*, p. 27 (May 7, 1990).
14. *Chemical Week*, **124** (1), 32 (1979).
15. "Ozone Depletion Accord," *Chemical and Engineering News*, p. 6 (July 9, 1990).
16. Bailey, F. E., Jr. and Koleske, J. V., *Alkylene Oxides and Their Polymers*, 261 pp., Marcel Dekker, Inc., New York, 1991.
17. Vandenberg, E. J., *J. Polymer Sci.*, A1, **7**, 525 (1969).
18. Davies, S. G. et al., *Tetrahedron*, **42**, 5123, (1986); *Tetrahedron Let.*, **27**, 3787 (1986).
19. Evans, D. A., *Science*, **240**, 420 (1988).
20. Schultz, P. G., *Science*, **240**, 426 (1988).
21. "The Next 75 Years," *Chemical Week*, p. 30 (Sept. 27, 1989).
22. "Stereolithography: An Introduction," *CHEMTECH*, p. 615 (Oct. 1990).
23. Hull, C. U., U.S. Patent 4,575,330 (1986).

2

Pollution Prevention and Waste Minimization

William J. Lacy, D.Sc., P.E., D.E.E.*

INTRODUCTION

Pollution prevention offers industry an opportunity, but its exact cost, benefits, and risks are difficult to fully identify or quantify. Pollution prevention represents a significant change in the scope and methodology usually used in waste management. It is a multimedia approach that concentrates on preventing the production of waste in any form in all parts of the plant.

Waste minimization was first introduced as a national policy in the 1984 Hazardous and Solid Waste Amendments (HSWA) to the Resource Conservation and Recovery Act (RCRA). The EPA elaborated on the concept in its 1986 Report to Congress.[1]

The general proposition is this: By practicing pollution prevention, industry can obtain improved environmental protection and increased industrial efficiency, profitability, and competitiveness. Preventing pollution at its source, through modifications in manufacturing processes or product design, is a goal.

There are numerous examples of successful waste-reduction initiatives and some impressive data, usually on a waste stream or plant basis. However, no comprehensive company, industry, or national data thus far available demonstrate broad success at cutting industrial-waste generation. Enormous amounts of wastes, pollutants, and discharges still continue to be generated.

It is not a matter of merely choosing between the traditional end-of-pipe treatment or pollution prevention. Preventing pollution is like preventing disease by changing eating habits and lifestyle; pollution control is like using medicine and surgery to minimize ill effects.

Pollution control can shift pollution around. Air and water pollution-control equipment extracts harmful substances and can generate large amounts of solid and hazardous waste which are disposed of in landfills. Some of these landfills have been linked to groundwater contamination. Regulatory gaps, regulatory noncompliance, the difficulty of responding to newly identified environmental problems, threats from very small residual levels of pollution, and continuing global population

*Formerly Principal Engineering-Science Advisor and Research Director at EPA Headquarters, Washington, D.C. Currently consultant to industry and government.

growth and industrialization all strongly support the goal of environmental protection.

Pollution prevention can help meet the public demand for solutions to urgent and extremely complex environmental problems, such as dealing with the threat of global warming and safely managing toxic and solid waste. This is why many environmentalists advocate pollution prevention as one possible way of addressing environmental problems.

Pollution prevention can also make good economic sense. The costs to government (taxpayers) of developing and enforcing pollution-control regulations and the costs to industry of complying with these regulations have become burdensome. All concerned need a better understanding of the technical means of pollution prevention and of the human, organizational, and social obstacles to applying it. Practicing pollution prevention may very well mean changing the way in which products are made.

To be effective in a given case, waste minimization must become organization or company policy. This policy must span the range from inventory management through equipment and process modification to reuse. Thus, waste minimization must become an integral part of the company's operations. Required actions would include:

- *Inventory management and improved operations*:
  Inventory and trace all raw materials.
  Purchase fewer toxic and more nontoxic production materials.
  Implement employee training and management feedback.
  Improve material receiving, storage, and handling practices.
- *Modification of equipment*:
  Install equipment that produces minimal or no waste.
  Modify equipment to enhance recovery or recycling options.
  Redesign equipment or production lines to produce less waste.
  Improve operating efficiency of equipment.
  Maintain a strict preventive maintenance program.
- *Production process changes*:
  Substitute nonhazardous for hazardous raw materials.
  Segregate wastes by type for recovery.
  Eliminate sources of leaks and spills.
  Separate hazardous from nonhazardous wastes.
  Redesign or reformulate end products to be less hazardous.
  Optimize reactions and raw material use.
- *Recycling and reuse*:
  Install closed-loop systems.
  Recycle on-site for reuse.
  Recycle off-site for reuse.
  Exchange wastes.

INCENTIVES

EPA has found that U.S. industry has the potential to significantly reduce public health and environmental risks by minimizing its hazardous waste production. Such reduction could also lower disposal costs and the nation's need for treatment and disposal facilities. Therefore, EPA is actively encouraging industry to find ways to reduce both the volume and the toxicity of its wastes. EPA is also working closely with the states to help foster waste reduction technologies.

For example, only 4 percent of the total amount of hazardous waste generated in 1981 was recycled, according to a recent Office of Technology Assessment (OTA) study.[2] Most industry efforts at waste minimization have resulted from efforts to decrease manufacturing costs, improve production yields, or comply with government standards. According to the study, federal and state regulations for disposal have increased land disposal costs from $15 per ton to a current cost of $250 per ton. Incineration costs have risen from $500 to $1500 per ton.[2]

WASTE MINIMIZATION TERMS

To ensure that readers of this chapter have a common understanding of the technical terms used, several key terms are defined below. Waste minimization techniques showing

Fig. 2.1. Waste minimization techniques.

source reduction and on-site/off-site recycling are presented in Fig. 2.1.

Waste minimization. The reduction of hazardous waste generated or subsequently treated, stored, or disposed of. This includes source reduction undertaken by a generator to reduce the total volume or quantity of hazardous waste and/or on-site and off-site recycling.

Source reduction. The reduction or elimination of waste generation at the source (usually within a process). This includes process changes, feedstock changes, improved housekeeping/management, and in-process recycle.

Waste reduction. In-plant practices that reduce, avoid, or eliminate the generation of hazardous waste. OTA states: "Actions taken away from the waste generating activity, including waste recycling or treatment of wastes after they are generated, are not considered waste reduction." Also: "Action that merely concentrates the hazardous content of a waste to reduce waste volume ... is not waste reduction." [3]

Recycling. Use or reuse of a waste as a substitute for a commercial product or as a feedstock to an industrial process. This includes on-site or off-site reclamation of useful fractions of a waste or removal of contaminants from a waste to allow its reuse.

Already, some U.S. companies have discovered how waste reduction can offer substantial benefits quickly at low cost. In the past, EPA's Industrial Pollution Control Program emphasized process modification to control industrial effluents.[4-6]

There are a variety of incentives for minimizing waste, including the following:

- Economic benefits, including reduction of the amount of waste to be treated and disposed of and reduction of raw materials costs.
- Reduced liability (including liability for both short- and long-term environmental problems, and for long-term monitoring of health and safety).
- Improved public image and demonstrated environmental concern.

The economics of waste minimization efforts are directly related to the costs of wastewater treatment and waste disposal. As treatment and disposal costs rise, there is added benefit to reducing the volume and toxicity of waste requiring treatment and disposal. Environmental regulations, especially those of RCRA and CWA (Clean Water Act) have had a major effect on treatment and disposal costs.[7]

DEFINITIONS USED BY EPA AND OTA

Table 2.1 shows EPA definitions of several pollution prevention terms and Table 2.2 shows the differences in definitions between the EPA's "Report to Congress: Minimization of Hazardous Waste" and the OTA's "Serious Reduction of Hazardous Waste Report."

ASSESSMENT PROCEDURE

To be able to identify waste minimization opportunities in a given facility or process, it is necessary to use a systematic approach that takes into consideration all of the important factors. The Waste Minimization Assessment Procedure (WMAP) represents such an approach. One part of a larger waste minimization program that is required of hazardous waste generators, the procedure involves a step-by-step approach to (1) understand the facility's wastes and processes, (2) identify options for reducing waste, and (3) determine which of the options exhibit sufficient technical and economic feasibility to justify implementation. Naturally, it is necessary to modify the procedure to fit the specific needs of individual companies. Thus, it should be viewed as a source of ideas and concepts, rather than a vigorous prescription of how to do assessments.

The waste minimization procedure, which is depicted in Fig. 2.2, consists of four phases:[7]

Planning and organization
Assessment
Feasibility analysis
Implementation

These phases are discussed below, and a sample assessment worksheet is presented.

Planning and Organization

Careful planning and organization are necessary to bring about a successful waste minimization program. To start the program and maintain momentum and control, it is necessary that there be strong management commitment from the very beginning. The program should set general goals by which to measure its effectiveness. Selecting a qualified program staff is critical to the ultimate success of the program.

Assessment

The assessment phase serves to identify the best options for minimizing waste through a thorough understanding of the waste-generating processes, waste streams, and operating procedures. Therefore, the initial tasks in the assessment phase include collecting information about the facility's waste streams, processes, and operations.

Information about the facility's waste streams can come from a variety of sources, such as hazardous waste manifests, biennial reports, environmental audits, emission inventories, waste assays, and permits. Mass balances should be developed for each of the important waste-generating operations to identify sources and gain a better understanding of the wastes' origins.

TABLE 2.1 Definitions of Pollution Prevention Terms*

Waste. In theory, the term "waste" applies to nonproduct outputs of processes and discarded products, irrespective of the environmental medium affected. In practice, since the passage of the Resource Conservation and Recovery Act (RCRA), most uses of the term "waste" refer exclusively to the hazardous and solid wastes regulated under RCRA, and do not include air emissions or water discharges regulated by the Clean Air Act or the Clean Water Act. The Toxics Release Inventory, TRI, refers to wastes that are hazardous as well as nonhazardous.

Pollution/Pollutants. In this report, the terms "pollution" and "pollutants" refer to all nonproduct outputs, irrespective of any recycling or treatment that may prevent or mitigate releases to the environment.

Waste Minimization. One of the earliest initiatives in pollution prevention was waste minimization. The initial focus here was on wastes (as defined under RCRA), rather than on a comprehensive evalution of industrial emissions regulated under all environmental statutes. This term become a source of controversy because some considered it to designate approaches to *treating* waste so as to minimize its volume or toxicity, rather than decreasing the quantity of waste at the source of its generation. The distinction became important because some advocates of decreased waste generation believed that an emphasis on waste minimization would deflect resources away from prevention towards treatment (the EPA Office of Solid Waste Source Reduction and Recycling Action Plan (see Chapter 5) will formally define waste minimization). In the current RCRA biennial report, waste minimization refers to source reduction and recycling activities, and now excludes treatment and energy recovery.

Industrial Source Reduction. Industrial source reduction is defined in the recently enacted Federal Pollution Prevention Act as "any practice which 1) reduces the amount of any hazardous substance, pollutant, or contaminant entering any waste [pollutant] stream or otherwise released into the environment (including fugitive emissions) prior to recycling, treatment, and disposal; and 2) reduces the hazards to public health and the environment associated with the release of such substances, pollutants, or contaminants. The term includes equipment or technology modifications, process or procedure modifications, reformulation or redesign of products, substitution of raw materials, and improvements in housekeeping, maintenance, training, or inventory control."

Source reduction does not entail any form of waste management (e.g., recycling and treatment). The Act excludes from the definition of source reduction "any practice which alters the physical, chemical, or biological characteristics or volume of a hazardous substance, pollutant, or contaminant through a process or activity which itself is not integral to and necessary for the production of a product or the providing of a service.

Waste Reduction. This term has been used by the Congressional Office of Technology Assessment and INFORM to mean source reduction. On the other hand, many different groups have used the term to refer to waste minimization. Therefore, care must be employed in determining which of these different concepts is implied when the term "waste reduction" is encountered.

Toxic Chemical Use Substitution. The term "toxic chemical use substitution" describe replacing toxic chemicals with less harmful chemicals, although relative toxicities may be fully known. Examples would include substituting a toxic solvent in an industrial process with a chemical with lower toxicity and reformulating a product so as to decrease the use of toxic raw materials or the generation of toxic byproducts. One such case involves DowBrands, a subsidiary of Dow Chemical Company, which responded to pressure from California public interest groups to reformulate its K2r* Spotlifter, excludes the toxin perchloroethylene.

In this report, this term also includes attempts to reduce or eliminate the use in commerce of chemicals associated with health or environmental risks. Examples include the phase-out of lead in gasoline, the attempt to phase out the use of asbestos, and efforts eliminate emissions of chlorofluorocarbons and halons. Some of these attempts have involved substitution of less hazardous chemicals for compariable uses, but other involve the elimination of a particular process or product from the market without direct substitution.

Toxics Use Reduction. The term "toxics use reduction" refers to the activities grouped under "source reduction," where the intent is to reduce, avoid, or eliminate the use of toxics in processes and/or products so as to reduce overall risks to the health of working consumers, and the environment without shifting risks between workers, consumers, and parts of the environment.

Industrial Pollution Prevention. The terms "industrial pollution prevention" and "pollution prevention" refer to the combination of industrial source reduction and toxic chemical use substitution. It does not include any recycling or treatment of pollutants, nor does it include substitution of a nontoxic product made with *nontoxic* chemicals for a nontoxic product made with *toxic* chemicals.

*Source: Pollution Prevention 1991, "Progress on Reducing Industrial Pollutants" EPA 21P-3003 (October 1991).

TABLE 2.2 Waste Covered by EPA and OTA Reports

EPA	*OTA*	*HSWA on waste minimization*
Hazardous Waste:		
EPA does not explicitly say which wastes are covered by its use of the term "hazard wastes." An assumption can be made that since the report deals within the context of RCRA and was mandated under RCRA that EPA considers waste minimization to cover only those solid wastes regulated as hazardous waste under RCRA.	"All nonproduct hazardous outputs into all environmental media, even though they may be within permitted or licensed limits. This is much broader than the legal definition of hazardous solid waste in the Resource Conversation and Recovery Act, its amendments, and subsequent regulations. Hazardous refers to harm to human health or the environment as is broader than the term "toxicity." For example, wastes that are hazardous because of their corrosive, flammability, explosiveness, or infectiousness are not normally considered toxic." [p. 31]	The phase used in HSWA was "hazardous wastes." The assumption can be made, therefore, that the waste minimization regulations required by HSWA were intended to cover only those solid wastes regulated as hazardous waste under RCRA. Some ambiguity exists, however, due to the legislative history. Senate Report No. 98-284 on waste minimization provisions in S. 757, voiced concerns about "... pollutants contained in effluents, emission's wastes or other pollution streams."
EPA	*EPA*	*OTA*
Differing views of multimedia approach:		
Waste minimization is RCRA: All incentives/disincentives (barriers) are framed within the RCRA content. The one exception is: "Commercial recycling facilities that wish to increase their operations might be reluctant to do so if the expansion were to require a revision of their NPDES water pollution permit to authorize a change in the composition of their discharges or allow for larger flows." [p. 29] Almost all information/data reviewed for report and assessment of needs for future concerns RCRA hazardous waste generation and management.	"To achieve its purpose, waste minimization like other pollution control measures, must look comprehensively across all environmental media, reduction in hazardous waste must not be made at the expense of increases in air or water pollution.... Waste minimization programs must therefore be carefully designed to avoid cross-media transfers and to protect human health and the environment in a comprehensive sense. The need to design a waste minimization program that addresses both of these goals provides a framework for integrating the objectives of all environmental programs." [p. 10]	*Waste reduction is multimedia*: "Reduction—applied to a broad universe of emissions, discharges, and waste—is the best means of achieving pollution prevention." [p. 7] "OTA has concluded that a comprehensive multimedia (air, water, land) definition for hazardous waste is necessary... 1) to avoid creating opportunities for shifting waste from one environmental medium to another possibly unregulated or less regulated medium ... and 2) to include wastes that are not currently regulated, such as most toxic air emissions. If the term hazardous waste is defined or applied narrowly, waste reduction measures can be ineffective." [p. 11]
Waste minimization is multimedia: EPA lists protecting human health and the environment as a key role for waste minimization because "... none of EPA's environmental control programs can fully eliminate all the risks that they attempt to control." [p. 9]	"EPA believed that the waste minimization must be implemented as a general policy throughout the hazardous waste management system and, ultimately, more broadly throughout all of EPA's pollution control programs." [p. 121]	

Sources: Compiled by OTA, 1987. For column 1, U.S. Environmental Protection Agency, Report to Congress: *Minimization of Hazardous Waste*, EPA/530-SW-86-033 (Washington, DC, EPA, Office of Solid Waste and Emergency Response, October 1986), pages as noted. For column 2, U.S. Congress Office of Technology Assessment, Serious Reduction of Hazardous Waste, OTA-ITE-317 (Washington, DC: U.S. Government Printing Office, September 1986), pages as noted.

Fig. 2.2. A waste minimization assessment procedure.

Collecting waste-stream data and constructing mass balances will create a basis by which the assessment can track the flow and characteristics of the waste streams over time. This will be useful in identifying trends in waste generation and will also be critical in the task of measuring the performance of implemented waste minimization options later. The result of the activity is a catalog of waste streams that provides a description of each waste, including quantities, frequency of discharge, composition, and cost of management.

In addition to data about waste streams, other information is needed to fully understand the facility's operations, including the following items:

- Process, equipment, and facility design information.
- Raw material production information.
- Operating cost information.
- Policy and organizational information.

Ideally, assessments should be conducted on all of the waste-generating operations in a plant. However, in larger plants this often

is not practical, considering the limited resources available. In these cases, a program task force should prioritize the streams. Important criteria to consider in prioritizing waste streams and/or facility areas include the following:

- Compliance with current environmental regulations.
- Disposal cost and/or volume of the waste.
- Hazardous nature of the waste, and other safety considerations.
- Potential for (and ease of) minimization.
- Potential for removing production or waste treatment bottlenecks.
- Available budget and expertise for the waste minimization assessment program.

A practical consideration in determining which waste streams to select is to find those that can be reduced with minimum economic or production impact. A successfully implemented waste minimization project will ensure the acceptance of further waste minimization efforts within the organization.[8]

The assessment team must include people who are familiar with the area of the facility to be assessed (e.g., first line operators and the production supervisors are recommended). These people may or may not already be on the assessment program task force. (In a large facility, the task force should have a broad understanding of the facility's operations, whereas the assessment team should have a specific understanding of the areas to be assessed.) It may be advisable to include people from other parts of the facility that regularly interact with the areas to be assessed.

Although collecting and reviewing data is important in the assessment, the assessment team must be familiar with the actual operation at the site. Therefore, the assessment team should visit the site during various stages or cycles of an operation. If all of the assessment team members work at the facility (or are located relatively close by), it is easy for the team members to visit the site. However, if some members are from outside the facility, it is recommended that a formal site inspection be carried out.

A formal inspection can address many of the questions raised by data collected earlier in the assessment phase. An inspection can also confirm whether the facility is operating in the way it was originally intended to operate. An inspection concentrates on understanding how the wastes are generated.

The assessment team should "walk the line" from the beginning of the process to the point where products and wastes leave the facility. Because waste can be generated in receiving and storage areas as well as the production areas, all areas within the site should be visited. The following guidelines will help in organizing an effective site inspection:

- Prepare an agenda in advance.
- Schedule the inspection to coincide with the particular operation of interest.
- Interview operators, foremen, and supervisors, and assess the operating personnel's awareness of the waste generation aspects of the operation.
- Observe the housekeeping aspects of the operation.
- Assess the overall cleanliness of the site.
- Review the organizational structure and the level of coordination of waste-related activities between the assessed facility area and other related areas.
- Assess the administrative controls.

After collecting data during the assessment preparation step and the site inspection, the members of the assessment team will have begun to identify possible ways of reducing waste in the assessed area. The generation of options is both a creative and an analytical process. Although individual assessment team members may be able to suggest many potential waste minimization options, the process can be enhanced by using group decision techniques such as brainstorming. These techniques allow the team to identify options that the individual members might not have come up with on their own.

Identifying potential options requires the expertise of the assessment team members, much of whose knowledge comes from their education and on-the-job experience. Other

sources of background information on potential options include the following:

- Trade associations.
- Published literature.
- Environmental conferences and exhibits.
- Equipment vendors.
- Plant personnel (especially the operators).
- Federal, state, and local government environmental agencies.
- Consultants and/or employees from other facilities.

A successful assessment will result in the proposal of many waste minimization options; so it is necessary to identify those options that offer a real possibility of minimizing waste and reducing costs. The purpose of the screening step is to eliminate those options that are perceived to be impractical. The screening of options may be achieved through an informal decision made by the assessment program manager, a vote of the assessment team, or a weighted sum method that combines relative weights of such factors as operating cost reduction, capital cost requirement, and reduction in the waste hazard.

Some options (such as procedural changes) may involve limited capital costs and may be implemented in a short time frame. The screening step should consider the ease of implementation. If such an option is clearly desirable and indicates a potential cost saving, it should be considered for further study.

In screening the options, the assessment team should address questions that relate to the waste minimization goals and constraints. Examples of questions that should be addressed include the following:

- Does the necessary technology exist to develop the option?
- How much will the option reduce the waste quantity, hazard, and treatment/disposal costs?
- How much will the option reduce the use of raw materials?
- What will the impact be on liability and insurance costs?
- How much does it cost? Is it cost-effective?
- Can the option be implemented within a reasonable amount of time?
- Does the option have a good "track record"? If not, is there evidence that the option can work in this case? What other benefits will occur?

Feasibility Analysis

The waste minimization options that are successfully screened in the assessment phase then undergo a more detailed feasibility analysis. Most important, the feasibility analysis consists of three steps: technical evaluation, economic evaluation, and implementing option selections.

The purpose of the technical evaluation is to ensure that the option will work as intended, given specific facility constraints and product requirements. Typical questions of concern in performing technical evaluation are:

- Will the option work in this application?
- How has it worked in similar applications?
- Is space available? Are utilities available? Must new utility systems be installed?
- Is the new equipment or procedure compatible with the facility's operating procedures, work flow, and production rates?
- How long will production be stopped in order to install the system?
- Will product quality be maintained, improved, or adversely affected?
- Is special expertise required to operate or maintain the new system? What service will the vendor provide? Does the system or procedure create or remove safety hazards?
- Does the system or procedure create other environmental problems?

All affected groups in the facility should contribute to and review the results of the technical evaluation. Prior consultation and review with the affected groups is needed to ensure the viability and the acceptance of the option. If the option calls for a change in

production methods, its effects on the quality of the final product must be determined. Only those projects that are judged to be feasible and practical in the technical evaluation go on to an economic evaluation.

An economic evaluation is carried out by using the standard measures of profitability, such as payback period or discounted cash flow techniques (internal rate of return and net present value). Each company uses its own economic evaluation procedures and criteria for selecting projects for implementation. In performing the economic evaluation, planners must consider various costs and savings. As in any project, cost elements can be grouped into capital costs and operating costs.

Capital costs for waste minimization projects include the fixed capital costs for designing, purchasing, and installing equipment as well as costs for working capital, permitting, training, start-up, and financing charges. It is important to realize that some waste minimization options, such as procedural or materials changes, will not have any capital costs. Also, many source reduction options have the advantage of not requiring environmental permitting to be implemented.

Waste minimization projects need to show a savings in operating costs to be economically effective. Operating costs and savings typically associated with waste minimization projects include the following:

- Reduced waste treatment, disposal, and reporting costs.
- Raw material cost savings.
- Insurance and liability savings.
- Increased costs (or savings) associated with product quality.
- Decreased (or increased) utilities, operating, and maintenance costs and overhead costs.
- Increased (or decreased) revenues from changes in production of marketable by-products.

Once the capital and operating cost savings have been determined, a project's profitability can be determined by using profitability measures. These profitability measures are discussed in virtually all financial management, cost accounting, and engineering economics textbooks. Those options that require limited capital costs should be implemented as soon as savings in operating costs can be shown.

An important consideration of waste minimization projects is their potential to reduce the toxicity of a waste stream and thereby to reduce the environmental risk. Although these risks can be qualitatively identified, it is difficult to quantify their financial impact. Also, although the profitability of a waste minimization assessment program is important in deciding whether to implement a project, compliance with environmental regulations must be considered. Violations may ultimately result in shutting down a facility, or carry possible criminal penalties for the company's responsible people.

The product of a feasibility analysis is a report that presents technical and economic information concerning the feasibility of each option. It also contains recommendations to implement the options. The report should include project cost and performance information. Topics covered in the report should include:

- Applicability of the technology.
- Resource requirements (resources available in-house and outside).
- Estimated construction period and production downtime.
- Means to evaluate performance after the project is implemented.
- Environmental and safety liability considerations.

Before the report is completed, the results should be reviewed with the appropriate individuals.

Implementation

Implementation of the waste minimization project requires management's commitment to overcome the natural resistance to change that occurs within an organization. Once the project is implemented and is operating, it is important to evaluate its performance. Is it

performing as expected? If not, should it be modified, or is its use still beneficial? What other potential options were identified during the implementation and operation of this option?

The waste minimization program should be viewed as a continuing one. As waste minimization options are implemented, the task force should continue to look for new opportunities, assess other waste streams, and consider options that were not pursued earlier. The ultimate goal is to reduce the toxicity and volume of wastes to the maximum extent practical.

TABLE 2.3 A Waste Minimization Worksheet

Number and Title	*Purpose/Remarks*
Planning and Organization	
1. Management Commitment	Needs top-down approval.
2. Assessment Overview	Summarizes the overall assessment procedure.
3. Program organization	Records key members in the WM program task force and the WM assessment teams. Also records the relevant organization.
4. Assessment Task Force Makeup	Lists names of assessment team members as well as duties. Includes a list of potential departments to consider when selecting the teams.
Assessment Phase	
5. Collects Facility Data	Lists background information about the facility, including location, products, and operations. This is a checklist of useful process information to look for before starting the assessment.
6. Prioritizes and Selects Targets	Identifies hazardous components, production rate, revenues, and other information about products.
7. Selects Personnel Teams	Records information about the personnel who work in the area to be assessed.
8. Collects and Reviews Data	Records input material information for a specific production or process area. This includes name, supplier, hazardous component or properties, cost, delivery and shelf-life information, and possible substitutes.
9a. Inspects Site and Characterizes Waste	Records source, hazard, generation rate, disposal cost, and method of treatment or disposal for each waste stream.
b. Prepares Waste Stream Summary	Summarizes all of the information collected for each waste stream. This sheet is also used to prioritize waste streams to assess.
10a. Generates Option	Records options proposed during brainstorming or nominal group technique sessions. Includes the rationale for proposing each option.
b. Develops Option Description	Describes and summarizes information about a proposed option. Also notes approval of promising options.
11. Screens and Selects Options	Makes evaluation by weighted sum method.
Feasibility Analysis	
12. Technical Evaluation	Provides detailed checklist for performing a technical evaluation of a WM option. This worksheet is divided into sections for equipment-related options, personnel/procedure-related options, and materials-related options.
13. Economic Evaluation	Provides detailed list of capital and operating cost information for use in the economic evaluation of an option.
Implementation	
14. Justifies Projects and Funding for Equipment and Procedures	Summarizes important tasks to be performed during the implementation of an option. This includes deliverable, responsible person, budget, and schedule.
15. Evaluates Performance	Records material balance information for evaluating the performance of an implemented option.

Assessment Worksheet

A sample worksheet used in making waste minimization assessments is shown in Table 2.3. Designed to clarify the assessment procedure, the worksheet includes the phase of each step in the procedure, its number and title, and a brief description of its purpose or other useful remarks.

EXAMPLES OF ECONOMIC BENEFITS[11]

A number of waste minimization efforts have been documented by EPA's Office of Solid Waste. This section presents brief descriptions of 12 cases of waste minimization initiatives taken verbatim from EPA/530-SW-87-026. For additional information readers are referred to that publication, entitled "Waste Minimization, Environmental Quality with Economic Benefits."

A Texas chemical manufacturer of adiponitrile, a nylon intermediate, recently developed a new process that improves product yield while reducing by 50 percent the amount of aqueous waste generated. The company's original process generated 800 gallons of wastewater per minute, along with nonchlorinated waste solvents that had to be incinerated. The new process enables the firm not only to reduce the amount of wastewater that must be treated, but also to burn the waste solvents in the company's powerhouse. Steam generated by the burning of the waste solvents is used in the manufacturing process, thereby saving the company more than $10 million per year in fuel oil.

A Pennsylvania die manufacturer uses 1,1,1-trichloroethane to clean and degrease machine parts. Prior to installing this solvent recovery unit, the company shipped the contaminated solvent offsite for reclamation and then purchased reclaimed solvent at $0.80 per gallon and virgin 1,1,1-trichloroethane at $4.50 per gallon. Using this solvent recovery unit, the company now reclaims solvent onsite at a cost of $0.04–$0.10 per gallon. In addition, the company's purchase of virgin 1,1,1-trichloroethane has dropped from two 55-gallon drums each month to two 55-gallon drums every 6 months, a savings of nearly $5,000 per year.

In 1986, EPA sponsored a waste minimization assessment at an electric arc furnace steel-making facility. The assessment team examined waste minimization options, including source reduction and resource recovery, for the company's corrosive and heavy metal wastes. The assessment revealed that calcium fluoride (fluorspar) in the sludge generated during the neutralization of the pickling line wastewater could be economically recovered. Previously, the company had disposed of the sludge and purchased 1,000 tons of fluorspar per year as flux material for the steel-making process. The waste minimization option identified by the assessment team will save the company $100,000 per year in costs avoided to purchase fluorspar, and a further $70,000 per year because of a 30 percent reduction in the volume of sludge to be disposed of.

An electronics plant installed an electrolytic metal recovery cell to recover copper from waste generated in the production of telephone switching equipment. The process produces a better quality copper deposit on the cell's cathode plates, where the copper collects in half-inch-thick sheets. The cell recovers 75 pounds of copper per week, which is sold for $0.50 per pound—a total of about $2,000 per year. The use of the cell also has eliminated 1 drum of sludge per week, saving an additional $4,000 per year.

Evaporative recovery systems can minimize the volume of waste from metal-plating baths and recycle plating solutions by recovering 90–95 percent of the plating solution lost through dragout. The operating cost of the recovery system is only $0.08 per gallon, while the dragout sludge hauling and disposal costs are close to $1.00 per gallon. With only 5–10 percent of the dragout requiring waste treatment, waste handling and disposal costs have been reduced significantly.

During standard equipment-cleaning operations, hospitals, universities, and research centers, as well as many small and medium-sized businesses, such as metal finishers and furniture manufacturers, generate small amounts of waste solvents. These waste solvents can be recycled for reuse in cleaning operations using small, commercially available recovery units. Depending on the commercial value and amount of solvent recovered, the pay-back time for recycling equipment can be as short as 1 year. Since

transportation costs can be very high, even businesses that use only low volumes of solvents may find it more economical to recycle their waste solvents onsite rather than ship the wastes offsite for recovery or disposal.

An electric company uses a water-based electrostatic paint system instead of a conventional organic solvent paint system. This has resulted in improved quality of application, decrease of downtime from 3 percent to 1 percent, reduction in the generation of aromatic waste solvent by 95 percent, reduction in paint sludge by 97 percent, and increase of efficiency with up to 95 percent recovery and reuse of paint. The new system reduced hazardous waste disposal costs and decreased personnel and maintenance costs by 40 percent.

Ion exchange metal-recovery units are used to remove heavy metals from aqueous residues generated by electroplating, metal-finishing, electronic manufacturing, and metal-refining processes. Ion exchange systems are commercially available, are relatively compact, and use little energy.

The chemical and pigments department of a major chemical company in Kentucky manufactures freon. Low-quality hydrochloric acid, generated as a manufacturing byproduct, was previously disposed of in injection wells. The company recently installed a new freon manufacturing process that produces high-quality hydrochloric acid. By installing the new process and building an additional acid storage facility, the company now is able to sell approximately 22 million pounds per year of acid that was previously discarded.

A polymer products operation in Arizona uses reverse osmosis to eliminate the discharge of cyanide-containing rinse water from one of the company's four plating units. The process, which concentrates the cyanide and separates it from the rinse water, reduces the environmental impact of the discharge and conserves valuable plating materials and water treatment chemicals.

One of the largest chemical manufacturers in the Netherlands uses waste segregation, removal of solvents in water solutions by distillation, and other source reduction measures to reduce the company's annual wastewater output by 80 percent.

In Sweden, a major pharmaceutical producer initiated a program to recycle approximately 10,000 tons of hazardous waste solvents per year through the company's onsite distillation plant, thereby reducing by 60 percent the amount of solvent waste that was shipped offsite for disposal.

INTERNATIONAL PERSPECTIVE

The waste reduction activities of national governments may be of interest to policymakers for two reasons. First, the choices made by other countries help them in understanding the range of policies available in addressing waste problems. Second, the expertise gained by nations with a relatively long experience in waste reduction can provide important historical trend data. Some Western European governments have actively encouraged waste reduction for many years. To the extent that their experience in waste reduction results in more efficient processes and increased productivity, U.S. firms can benefit from Western Europe's experience. Currently, the competition among waste-reduction vendors favors Western Europe. If a worldwide market for waste-reducing technologies and techniques emerges in the coming decade, U.S. firms may find that European vendors are offering a wide variety of technologies that have been tested over a long period of time.

Some of the earlier initiatives in waste reduction came from international organizations. The United Nations Economic Commission for Europe (ECE) sponsored the first International Conference on Non-Waste Technology in Paris in 1976. In 1979, the ECE adopted a detailed "Declaration on Low and Non-Waste Technology and Reutilization and Recycling of Wastes."[12] In this document, the ECE recommended action on both the national and international levels to develop and promote low and non-waste technologies. International ECE activities resulting from this declaration included:

- Publication of a four-volume compendium on low and non-waste technologies in

	JAPAN	CANADA	GERMANY	SWEDEN	NETHER-LANDS	DENMARK
Tax Incentives						
Waste End Taxes			●		●	●
Tax Incentives	●	●	●			
ECONOMICS						
Price Support System for Recycling			●		●	
Government Grants as Subsidies	●	●	●	●	●	●
Low Interest Loans	●		●			
TECHNICAL ASSISTANCE						
Information and Referral Service	●	●	●		●	●
Site Consultation	●	●				
Training Seminars		●	●			●
R&D ASSISTANCE						
Technical Development Labs				●	●	●
Demonstration Projects	●	●	●	●		
Industrial Research			●	●	●	
PERMITS AND PLANS						
National Waste Managements Plans					●	
Waste Reduction Agreements	●					
Waste Reduction as a Part of Permits				●		
WASTE EXCHANGE						
Regional Waste Exchanges	●	●	●		●	●
PUBLIC INFORMATION						
Focus on Corporate Image	●					
Focus on Consumer Practices			●			

Fig. 2.3. Waste minimization practices in other countries. (*"Foreign Practices in Hazardous Waste Minimization," Medford, MA: Center for Environmental Management, Tufts University, 1986.*)

1982, listing over 80 examples of successful pollution prevention efforts by European industrial firms.[12]

- Publication of a compendium of lectures by experts in low and non-waste technology in 1983.[13]
- Holding a European Seminar on Clean Technologies at The Hague in 1980.
- Setting up a Working Party on Low and Non-Waste Technology and Re-utilization and Recycling of Wastes, which has met annually since 1980.[14]
- Setting up an environmental fund for demonstrating innovative technologies that are broadly applicable to reducing pollution. A sum of 6.5 million in European Currency Units (about 6.1 million in U.S. dollars) was set aside for this purpose in 1985.[15]
- Holding an International Solid Waste Association meeting in Geneva, Switzerland in 1989.[16]

Various waste minimization practices used in other countries are shown in Fig. 2.3.

IMPLEMENTATION OPTIONS AND SOURCES OF HELP

After a waste minimization program is implemented, the question is, "How effective are the options?" Means to measure the effectiveness need to be in place. Reduction could be achieved in the volume of waste produced or in the degree of hazard. Another measure could be the ratio of waste quantity to material throughput or product output.[17] This measure is applicable for individual

operations or processes. Another option could be implemented at other facilities that utilize the same process or operation within the plant.[18]

As pointed out by Newton,[19] waste minimization must not be a one-time effort. It should be continuous. Areas with lower priorities must be addressed once the higher priorities have been implemented. Assessments should be conducted yearly or when there are:

- Changes in raw material or product requirements.
- Changes in waste management costs.
- New technologies available.
- Changes in regulations.
- Environmental events.

Additional information on waste reduction is available from a wide variety of source materials, including:[20–22]

- Sector or specialized trade associations.
- Consulting/engineering firms that offer assistance in identifying and classifying waste, drum handling, site assessment, and preparation of manifests.
- Hazardous waste management contractors who offer information about costs involved with treatment, disposal, and waste transport.
- Chemical suppliers who offer information on chemical safety and waste management.

Potential opportunities for waste reduction in several industrial areas are listed in Table 2.4.

In conclusion, pollution prevention is one aspect of environmental protection. Prevention

TABLE 2.4 Potential for Waste Reduction Opportunities across Different Industry Types

Company/Industrial Characteristic	*Example Industries*	*Operations Changes*	*In-Process Recycling*	*Process Changes*	*Input Substitution*	*End Product Changes*
Mature process technology, high volume product	Rubber Petroleum Commodity chemicals Paper products Lumber	+	+	–	–	–
Very stringent product specifications or high product quality demands for high cost/high profit products	Pharmaceuticals Weapons Robotics Specialty chemicals	+	–	–	–	–
Frequently changing, high-tech products for industrial use	Electronic components Medical equipment	+	+	+	+	–
Job shop processing of many different industrial products	Electroplating, Printing	+	+	+	+	–
	Foundries, Machine shops	+	+	–	+	–
Changing production technology for commodity goods	Steelmaking Nonferrous metals Textiles	+	+	+	–	–
Large-scale manufacture of consumer goods	Automobiles Appliances Consumer electronics Paints	+	+	+	+	+

Source: Office of Technology Assessment, J. Hirschhorn (1986).

of pollution at its source by modification of processes, equipment, product design, or utilization of alternative raw materials is the goal. Pollution prevention is a broad term that covers energy conservation, recycle, and reuse, whereas waste minimization, the major theme of this chapter, concentrates on waste streams. The ultimate goal of waste minimization is to reduce the generation of waste to the greatest extent possible.

REFERENCES

1. U.S. Environmental Protection Agency, "Report to Congress: Minimization of hazardous Waste," EPA/530-SW-86-033, OSW and ER, Washington, DC (1986).
2. U.S. Congress, Office of Technology Assessment, "Serious Reduction of Hazardous Waste," OTA-ITE-317, GPO, Washington, DC (Sept. 1986).
3. Hirschhorn, Joel S., and Oldenburg, Kirsten V., "Taking Seriously the Reduction of Hazardous Waste" *Pollution Engineering* (Jan. 1987).
4. Lacy, W. J., and Keeler, H. G., "Process Modification to Abate Pollution in the Food Industry," presented at the International Union of Pure and Applied Chemistry Conference held in Stockholm, Sweden, Nov. 1974.
5. Lacy, W. J., Rey, G., and Keeler, H. G., "Process Modification for Industrial Water of Pollution Control," presented at the meeting of the American Institute of Chemical Engineers, Chicago, IL, Aug. 1983.
6. Lacy, W. J., address to the Plenary Session of USEPA-ASTM-High Public Health Research Institute, 3rd International Conference (*Proceedings* published by ASTM), Alexandria, Egypt, June 1985.
7. Lorton, Gregory A., Fromm, Carl H., and Freeman, Harry, "The EPA Manual for Waste Minimization Opportunity Assessment," EPA/600/S2-88/02S (Aug. 1988).
8. EPA-ORD Hazardous Waste Engineering Research Lab., "Waste Minimization Opportunity Assessment Manual," EPA/62S/5-88/003 (July 1988).
9. Hirschhorn, Joel S., Oldenburg, Kirsten V., Buyrn, Audrey, and Johns, Lionel S., Office of Technology Assessment, "From Pollution to Prevention: A Progress Report on Waste Reduction," OTA-ITE-347, Washington, DC, GPO (June 1987).
10. Hirschhorn, Joel S., Oldenburg, Kirsten V., Buyrn, Audrey, and Johns, Lionel S., Office of Technology Assessment, "Serious Reduction of Hazardous Waste," OTA-ITE-317, Washington, DC, USGPO (Sept. 1986).
11. Alford, M., Schultz, H., Wilkes, A., Dellinger, R., and Schecter, R., "Waste Minimization, Environmental Quality with Economic Benefits," EPA/530-SW-87-026, OSW and ER Washington, DC (Oct. 1987).
12. U.N., ECE, "Declaration on Low and Non-Waste Technology of Reutilization and Recycling of Wastes," Geneva, Switzerland (Nov. 1979).
13. Hungarian National Authority for Environmental Protection and Nature Conservation, "Compendium of Lectures on Low and Non-Waste Technology," Budapest, Hungary (Dec. 1983).
14. O.E.C.D., "Resolution of the Council on International Cooperation Concerning Transfrontier Movements of Hazardous Wastes" (July 1985).
15. European Council of Chemical Manufactures Federations, "Industrial Waste Management," CEFIC, Brussels, Belgium (1985). As cited in Royal Commission on Environmental Pollution, 11th Report, "Managing Waste A Duty of Care," Her Majesty's Stationary Office, London 1985).
16. Jakobson, Kurt, "EPA's Waste Minimization Assessment Procedure: An Overview," presented at International Solid Waste Association Meeting, Geneva, Switzerland, May 19–June 1, 1989.
17. Warner, David J., and Cece, John M., Jr., "Waste Minimization Standards, A Significant Economic Incentive," *ASTM Standardization News* (Apr. 1987).
18. Hirschhorn, J. S., Rams, J. M., and Lacy, W. J., "Company Champions for Waste Reduction," *Chemistry for Protection of the Environment 1987*, Elsevier Environmental Sciences #34, pp. 399–410, Elsevier, New York, 1988.
19. Newton, Jim, "Setting Up A Waste Minimization Program," *Pollution Engineering*, pp. 75–80 (Apr. 1990).
20. "Waste Minimization Case Studies," *The Waste Consultant*, pp. 1-19–1-25 (Nov.–Dec. 1989).
21. Comella, Paula, and Rittmeyer, Robert W., "Waste Minimization/Pollution Prevention," *Pollution Engineering*, pp. 71–74 (Apr. 1990).
22. Lacy, W. J., and Rams, J. M., "Pollution Prevention Case Studies," presented at First International Conference of Pacific Rim Nations, Honolulu, Hawaii, Jan. 1988.

3

Industrial Wastewater and Hazardous Materials Treatment Technology

William J. Lacy, Dr. Sc., P.E., D.E.E.*

INTRODUCTION

The gravest water quality issue now facing the nation is the disposal of industrial wastes. Other major environmental problems include medical wastes, hazardous wastes, and toxic contamination of the nation's streams and groundwater. Municipal waste disposal and landfill issues also are very serious concerns.

It has been reported (*Environmental Business Journal*, July 1989) that:

- About 13,600 hazardous substance spills occur outside fixed facilities each year; OSHA reports that an additional 11,000 such spills occur inside the workplace.
- Some $13.1 billion of revenue was generated by firms providing hazardous waste disposal equipment and services in 1987; the figure is expected to double by the year 2000.
- By 1990 there were about 5666 facilities for hazardous waste treatment, storage, disposal, and recycling facilities operating in the United States. The Government Accounting Office estimated that some 2500 facilities will require some corrective action under the Resource Conservation and Recovery Act (RCRA).
- In 1987 9.7 billion pounds of toxic substances were released into streams and other surface waters. In the same year 1.9 billion pounds went to municipal wastewater treatment plants and 2.7 billion pounds to landfills, and 2.7 billion pounds of chemicals were emitted into the air. These figures do not include releases by small quantity generators.

The same journal also reported that:

- Current annual investment in pollution control equipment amounted to $200 billion.
- The estimated cost to clean up hazardous and radioactive waste at 45 DOE sites was $92 billion.
- The market for drinking water purification equipment and services in 1995 is expected to be $7.7 billion.

*Formerly Principal Engineering-Science Advisor and Research Director at EPA Headquarters, Washington, D.C. Currently consultant to industry and government.

- Some 1.2 billion gallons of waste oil are generated each year at gasoline stations, auto dealerships, and industrial facilities.

Figure 3.1 shows estimated market values and predicted annual growth rates for various environmental sectors.

Industrial effluents contain enormous quantities of inorganic and organic chemical wastes, which are steadily becoming more complex and difficult to treat by conventional technologies. The variety and the complexity of inorganic and organic compounds contained in these effluents present a serious liquid wastewater treatment control problem, in that the pollution and toxicity effects of these constituents are of much greater significance than those found in domestic wastewater.

Conventional wastewater treatment technologies, often barely adequate for existing waste types, offer even less promise of providing the type and the degree of treatment that will be required in the near future. Therefore, industrial pollution-control technology must be developed to achieve effective and economical control of pollution from such varied industries as those producing metals and metal products, chemicals and allied products, food and kindred products, textiles, and leather goods.

Fig. 3.1. Estimated environmental market values and, in parentheses, estimated annual growth rates for various environmental sectors. Market values do not include remediation. (*Data from R. K. Miller and Assoc.*, Waste Technology News, *Vol. 2, No. 22, July 1990.*)

Federal Laws and Regulations

There are currently 15 key federal laws related to the control and regulation of hazardous and toxic materials. A summary of these laws, their regulatory objectives, and primary responsible federal agencies is shown in Table 3.1. In addition there are many more state laws and regulations, often with more stringent requirements than their federal counterparts. It is important to review these state and federal regulations as part of the risk assessment process. The various laws and regulations provide specific definitions, specific lists of toxic and hazardous materials, requirements for compliance, and methods and standards for compliance.

TREATMENT LEVELS

Presently, there are two basic methods of treating wastewaters; in certain cases additional treatment is required. The processes or methods are referred to as primary, secondary, and tertiary treatment.*

Primary Treatment

As wastewater enters a plant for primary treatment, it flows through a screen that removes large floating objects such as rags and sticks that may clog pumps and pipes. Such screens vary from coarse to fine—from those consisting of parallel steel or iron bars with openings of about a half inch or more to screens with much smaller openings.

Screens are generally placed in a chamber or channel in a position inclined with respect to the flow of the sewage to make cleaning easier. The debris caught on the upstream surface of the screen can be raked off manually

*The reader is encouraged to consult the glossary of wastewater-treatment terms at the end of this chapter as needed.

TABLE 3.1 Summary of Key Federal Laws and Responsible Agencies Related to the Control and Regulation of Hazardous and Toxic Materials

Regulatory Objectives	*Key Federal Laws*	*Responsible Agencies*
Air pollution control	Clean Air Act (CAA)	EPA
Water pollution control	Clean Water Act (CWA)	EPA
	Safe Drinking Water Act	EPA
	Marine Protection, Research and Sanctuaries Act (MPRSA)	EPA, Corps of Engineers
Regulation of hazardous waste	Resource Conservation and Recovery Act (RCRA)	EPA, DOT
	Comprehensive Environmental Response, Compensation and Liability Act (CERCLA)	EPA
	Hazardous Materials Transportation Act	DOT, EPA
Regulation of the transportation of hazardous materials	Hazardous Materials Transportation Act	DOT
	Dangerous Cargo Act	DOT
	Federal Railroad Safety Act	DOT
	Ports and Waterways Safety Act	DOT
Regulation of the workplace	Occupational Safety and Health Act	OSHA, NIOSH
	Federal Mine Safety and Health Act	MSHA, NIOSH
Regulation of toxic industrial chemicals	Toxic Substances Control Act (TSCA)	EPA
Regulation of pesticides	Federal Insecticide, Fungicide and Rodenticide Act (FIFRA)	EPA
	Federal Environmental Pesticide Control Act	EPA
Update cleanup enforcement activities	Superfund Amendments and Reauthorization Act (SARA)	EPA

Note:
EPA—U.S. Environmental Protection Agency
DOT—U.S. Department of Transportation
OSHA—Occupational Safety and Health Administration
NIOSH—National Institute of Occupational Safety and Health
MSHA—Mine Safety and Health Administration

or mechanically. Some plants use a device known as a comminutor, which combines the functions of a screen and a grinder. These devices catch and then cut or shred the heavy solid material. The pulverized material remains in the sewage flow to be removed later in a settling tank.

After the wastewater has been screened, it passes into a grit chamber, where sand, grit, cinders, and small stones are allowed to settle to the bottom. The unwanted grit or gravel from this process is usually disposed of by land filling near a treatment plant.

In some plants, another screen is placed after the grit chamber to remove any additional material that might damage equipment or interfere with later processes.

With the screening completed and the grit removed, the wastewater still contains suspended solids. These are minute particles of matter that can be removed from the sewage by treatment in a sedimentation tank. When the speed of the flow of wastewater through one of these tanks is reduced, the suspended solids will gradually sink to the bottom. This mass of solids is called raw sludge.

Various methods have been devised for removing sludge from the sedimentation tanks. In older plants it was removed by hand. After a tank had been in service for several days or weeks, the sewage flow was diverted to another tank, and the sludge in the bottom of the out-of-service tank was pushed or flushed with water to a nearby pit and then removed for further treatment or disposal.

Almost all plants built within the past 30 years have included mechanical means for removing the sludge from sedimentation

tanks. In some plants it is removed continuously, whereas in others it is at intervals. To complete the primary treatment, the sludge-free effluent may be chlorinated to kill harmful bacteria and then discharged into a stream or river, or it can be sent to a secondary treatment process.

Secondary Treatment

Secondary treatment removes 90-plus percent of the organic matter in wastewater by making use of the bacteria it contains. The two principal processes for secondary treatment are trickling filters and the activated-sludge process. The effluent from the sedimentation tank in the primary treatment stage flows or is pumped to a facility using one or the other of these processes.

Trickling Filter. A trickling filter is simply a bed of stones from 3 to 10 feet deep through which the wastewater passes. Bacteria gather and multiply on these stones until they can consume most of the organic matter in the wastewater. The cleaner water trickles out through pipes in the bottom of the filter for further treatment. The wastewater is applied to the bed of stones in two principal ways. One method consists of distributing the effluent intermittently through a network of pipes laid on or beneath the surface of the stones. Attached to these pipes are smaller, vertical pipes that spray the effluent over the

Fig. 3.2. Types of secondary biological treatment methods. (*Courtesy EPA.*)

stones. Another much-used method consists of a vertical pipe in the center of the filter connected to rotating horizontal pipes that spray the effluent continuously upon the stones.

Activated-Sludge Process. The trend today is toward the use of the activated-sludge process instead of trickling filters. The former process speeds up the work of the bacteria by bringing air and sludge heavily laden with bacteria into close contact with the wastewater. (See Fig. 3.2.)

In the activated-sludge process, the wastewater from the sedimentation tank in primary treatment is pumped to an aeration tank where it is mixed with air and sludge loaded with bacteria and allowed to remain for

Fig. 3.3. Oxygen and contact stabilization processes for wastewater treatment. (*Courtesy EPA.*)

Fig. 3.4. Other types of aeration and biological treatment processes. (*Courtesy EPA.*)

several hours. During this time, the bacteria break down the organic matter. From the aeration tank, the wastewater, now called mixed liquor, flows to another sedimentation tank to remove the solids. Chlorination of the effluent completes the basic secondary treatment. The sludge, now activated with additional millions of bacteria and other tiny organisms, can be used again by returning it to an aeration tank for mixing with new wastewater and ample amounts of air.

The activated-sludge process, like most other techniques, has advantages and limitations. The size of the units needed is small so that they require comparatively little land space. Also, the process is free of flies and

odors. But it is more costly to operate than the trickling filter, and it sometimes loses its effectiveness when presented with difficult industrial wastes.

An adequate supply of oxygen is necessary for the activated-sludge process to be effective. Air is mixed with wastewater and biologically active sludge in order for the aeration process to be effective. This mixing can be done using three different methods. The first, mechanical aeration, is accomplished by drawing the wastewater from the bottom of the aeration tank and spraying it over the surface, thus causing the wastewater to absorb large amounts of oxygen from the atmosphere. In the second method, large amounts of air under pressure are piped down into the wastewater and forced out through openings in the pipe. The third method is a combination of mechanical aeration and the forced-air method. (See Figs. 3.3 and 3.4.)

Tertiary Treatment

Tertiary treatment is used when the waste stream must meet strict requirements governing recreational bodies of water, or must approach drinking-water standards. This may require one or several of the following processes: slow filtration; rapid filtration with activated carbon; adsorption by activated carbon; application of ozone; high-rate chlorination or the use of another oxidizing chemical; or lagooning.

At each plant the question may arise of what degree of treatment is actually required. Water-quality criteria imposed by different waste-streams discharges may vary widely. Even within the same state, or for a particular river basin, different limits for each of the contaminants may be set for the section of the river under consideration.

WASTE TREATMENT TECHNOLOGIES

Anaerobic Digestion

Anaerobic digestion is a biodegradation process capable of handling high-strength aqueous waste streams that would not be efficiently treated by aerobic biodegradation processes. Advantages of anaerobic systems over aerobic systems include:

- Ability to break down some halogenated organics.
- Low production of biomass sludges that require further treatment and disposal.
- Low cost.

However, anaerobic systems can be less reliable than aerobic systems. For this reason, aerobic systems are better suited for mobile (truck- or trailer-mounted) unit applications. Disadvantages of anaerobic systems include:

- Potential for shock loading of biomass and termination of biodegradation process due to variation in waste stream characteristics.
- Low throughput due to the slow biodegradation process (two steps).
- Frequent need for further treatment of effluent prior to discharge off-site or to a municipal treatment system.
- Generation of methane gas (a problem if it cannot be readily used on-site for meeting energy requirements).

Careful design and control often can solve these problems, but vendors are reluctant to recommend anaerobic mobile systems. Anaerobic systems are much more susceptible to variation in waste stream characteristics and environmental parameters. Fixed anaerobic systems are widely used in industry for treatment of uniform, concentrated biodegradable waste in aqueous waste streams due to the low-cost, low-residual generation and production of usable methane gas. However, application to variable CERCLA waste streams is relatively infrequent. Anaerobic systems have the best potential as a pretreatment step for an aerobic system that would otherwise be unable to process a high-strength waste such as a leachate.

Figure 3.5 shows schematic diagrams of conventional and high-rate anaerobic digesters.

Fig. 3.5. Schematic diagrams of conventional and high-rate anaerobic digesters. (*Courtesy EPA.*)

Waste Types Handled. Anaerobic digesters can handle concentrated waste streams with biodegradable nonhalogenated organics and moderate levels of halogenated organics. The most suitable application may be as a treatment step for landfill leachates where storage, mixing, and flow regulation can be accomplished prior to introduction at the digesters.

Anaerobic digestion can partially break down some halogenated organics unsuitable for aerobic digestion. Also, anaerobic systems can be used as a pretreatment step prior to aerobic biodegradation. Recommended influent characteristics for anaerobic processes are as follows:

BOD: 1000 to 15,000 mg/l
COD: >1500 mg/l

Restrictive Waste Characteristics. As with aerobic systems, the biodegradation process can be slowed or halted by the following:

- Abrupt change in waste stream characteristics.
- Variable environmental conditions (e.g., temperature, pH).
- Elevated levels of heavy metals or halogenated organics toxic to the biomass.
- Inadequate nutrient levels.

Wet Air Oxidation

Wet air oxidation is a thermal treatment technology that breaks down organic materials by oxidation in a high-temperature, high-pressure aqueous environment and in the presence of compressed air. The resulting exothermic reactions are self-sustaining and potentially capable of generating steam as a by-product.

In this process, wastes are mixed with compressed air. The waste–air mixture then is preheated in a heat exchanger before entering a corrosion-resistant reactor where exothermic reactions increase the temperature to the desired level. The exit stream from the reactor is used at the place where the spent process vapors (i.e., noncondensable gases consisting primarily of air and carbon dioxide) are separated from the oxygenated liquid phase.

The reactor or pressure vessel is sized to accommodate a specific waste flow over a certain amount of time. The residence time, temperature, pressure, and possibly a catalyst are based on the characteristics of the waste.

Fig. 3.6. Process flow diagram of wet air oxidation. (*Courtesy Zimpro Passavant Environmental Systems, Inc., Rothschild WI.*)

A process flow diagram for wet air oxidation is shown in Fig. 3.6.

Waste Types Handled. Wet air oxidation can be applied to dissolved or suspended organic substances in the form of liquid wastes and sludges. Particular contaminants and wastes processed include: halogenated organics, inorganic/organic sludges, contaminated groundwater, inorganic and organic cyanides, phenols, and leachates.

Although wet air oxidation technology has been used extensively for industrial applications, its use for hazardous wastes is limited. There are, however, several full-scale fixed wet air oxidation facilities treating hazardous waste.

Restrictive Waste Characteristics. Non-pumpable aqueous wastes and highly halogenated wastes are restricted from this process. Minimum and maximum chemical oxygen demand (COD) concentrations for a feasible application are 10,000 mg/l and 200,000 mg/l, respectively.

Supercritical Water Oxidation

The supercritical water oxidation process is based on the ability of water to perform as an excellent solvent for organics when the water is above its critical temperature (705°F) and critical pressure (3200 psi). When air is mixed with aqueous wastes above the critical temperature and pressure of water, organics are reported to be rapidly and completely oxidized to CO_2 and water. In addition, inorganic salts become almost insoluble above 930°F and precipitate out of the supercritical liquid. The exothermic conditions during the oxidation reactions produce energy in excess of process energy requirements and, in principle, allow for the production of high pressure steam or electricity.

Process operation involves the use of a high pressure pump to bring an aqueous solution or slurry of hazardous wastes up to system pressure before being heated to supercritical conditions in a feed/effluent heat exchanger. Large organic molecules are thus broken down to molecules of low molecular weight. High pressure air then is injected into the reactor, rapidly oxidizing the lower molecular weight compounds. Bases such as sodium hydroxide are added to the waste to neutralize any inorganic acids formed during oxidation.

The supercritical water oxidation process can be adapted to a wide range of feed mixtures and scales of operation, making it a

Fig. 3.7. Process flow diagram of supercritical water oxidation process. (*Modar, Inc., Houston, TX; courtesy EPA.*)

potentially feasible mobile technology. However, supercritical water oxidation is a relatively new thermal technology. To date, operating experience has been restricted to bench-scale and pilot-scale systems.

A process flow diagram is shown in Fig. 3.7.

Waste Types Handled. This technology can be applied to aqueous solutions or slurries with organic concentrations from 1 to 100 percent. The actual organic concentration of the waste fed to the process will depend on the heating value of the original waste material. The heat content of waste fed to the process is controlled at 1800 Btu/lb. Therefore, wastes with a heating value below 1800 Btu/lb require the addition of auxiliary fuel, and waste material with a heating value above 1800 Btu/lb requires either the addition of dilution water or blending with a lower-heating-value waste.

Particular contaminants and wastes processed include: PCBs, dioxins, solvents, pesticides, and still bottoms and tank bottoms.

Filtration

Filtration may be used for two primary purposes:

1. Removal of suspended solids from a fluid by passage of the fluid through a bed of granular material.
2. Dewatering of sludges and solids by vacuum, high pressure, or gravity. Granular media filters (typically sand and anthracite) remove suspended solids through straining, physical adsorption, and coagulation–flocculation. These filters may be pressurized or gravity-fed, with loading rates ranging from 2 to 15 gpm/sq ft.

Various filtration methods have been employed to dewater sludges, including:

- Vacuum filtration
- Belt filter press
- Chamber pressure filtration

Vacuum filtration typically is performed using a mechanically supported cylindrical rotating drum covered by a filter medium (cloth, coil springs, or wire mesh fabric). Water is drawn into the center by vacuum while the solids are scraped off the drum.

The belt filter press continuously squeezes the sludge through a series of rollers, which apply increasing pressure and shear force on the sludge.

The chamber pressure filters consist of a collection of cloth-covered plates arranged in parallel and pressed together by pressure up to 200 psi. As the plates are compressed, filtrate exits through the cloth.

Gravity-fed or pressurized granular media filter systems are less energy-intensive than the three systems mentioned above but require highly qualified operators with sufficient experience in backflushing procedures. These systems are not applicable to the removal of dissolved chemical species.

Waste Type Handled. Pressurized and gravity-fed granular media filtration systems are used for waste streams containing suspended solids. Vacuum, belt press, and pressure chamber filtration processes are used primarily to dewater sludges.

Restrictive Waste Characteristics. Energy-intensive filtration operations such as belt press filtration, vacuum rotary filtration, and pressure filters operate at optimal solids content. Sludges that range from 5 to 10 percent solids are generally well suited for these operations.

Ion Exchange

Anions and cations dissolved in a dilute aqueous waste can be removed from solution through the process of ion exchange. As the name implies, one ion, electrostatically attached to a solid resin material, is exchanged for a dissolved toxic ion. The exchange reaction is reversible, allowing for resin regeneration.

The exchange occurs because the divalent and trivalent toxic metal anions or cations have an increased affinity for the charged sites on the surface of the resins, which originally were coated with weakly held monovalent anions or cations such as chloride, hydroxyl, sodium, or hydrogen ions.

Currently, the majority of new ion exchange resins are constructed of synthetic organic materials. These resins are able to withstand a wide range of temperatures and pH, are capable of specific selectivity and a high exchange capacity.

It is possible to remove both dissolved toxic anions and cations by placing a cation exchange column and an anion exchange column in series. This system has the ability, depending upon the choice of resins, to remove a wide range of inorganic and organic dissolved contaminants.

Pressure vessels typically range from 2- to 6-inch-diameter systems up to a custom size of 12 feet in diameter. Corresponding flow rates range from 25 gpm up to a maximum of 1150 gpm. These vessels could easily be truck-mounted and moved from site to site.

Waste Types Handled. Wastes that are suited for anion exchange include:

- All metallic anions and cations such as $Cr_2O_7(-2)$, $SeO_4(-2)$, $AsO_4(-2)$, $Ni(+2)$, $Cd(+2)$, and $Hg(+2)$.
- Inorganic anions such as halides, sulfates, and cyanides.
- Organic acids such as carboxylics, sulfonics, and some phenols.
- Organic bases such as amines.

Restrictive Waste Characteristics. The upper limits of concentration at which ion exchange may be applied are 2500 mg/l for dissolved ions and 50 mg/l for suspended solids. Higher concentration levels of dissolved ion will result in rapid exhaustion of the resin with unusually

high regeneration expenses. High concentrations of suspended solids will result in clogged columns. Oxidants in the waste stream also should be avoided.

Membrane Separation

Membrane technologies separate solutes or contaminants from liquids through the use of semipermeable membranes, which function by selectively rejecting contaminants based on pore size, ionic valence, or co-precipitation. Membrane separation processes can be used for volume reduction, purification, concentration, and recovery of the contaminants.

The types of membrane separation technologies include reverse osmosis, hyperfiltration, ultrafiltration, and electrodialysis. At present, reverse osmosis is the only membrane separation technology that has been used as a mobile system and thus is the only such technology discussed in this section.

The reverse osmosis system allows a solvent (such as water) to be removed from a solution containing solutes (such as salts) by use of a pressure-driven membrane process. In this process, the solvent molecules (water) are forced through the microscopic pores of a semipermeable membrane by achieving sufficient hydrostatic pressure to overcome the concentrate osmotic pressure. Operating pressures for a typical reverse osmosis system can range from 200 to 800 psi. As the solvent flows through the membrane, the larger organic and inorganic compounds are rejected.

The effectiveness of RO systems depends on the following factors:

- Temperature of the solution
- Pressure of the system
- pH of the solution
- Chemical and physical structure of the membrane

Waste Types Handled. The reverse osmosis process has typically been used for treatment of brackish waters and aqueous metal wastes (plating baths), but innovative technologies have made it very effective in treating other forms of contaminants such as:

- PCB and chlorinated organics in groundwater
- Waters containing high BOD levels
- Insecticides/herbicides in groundwater
- Organic and inorganic leachate

Concentration levels of organics generally range in the milligram per liter level for the feed, with a permeate in the range of 10 to 50 μg/l.

Reactive Waste Characteristics. To avoid membrane plugging and to ensure a long maintenance-free life, it is important to remove suspended solids and oils with conventional pretreatment.

Mobile Incineration System (MIS)

Development of the EPA Mobile Incineration System began in the mid-1970s as a research effort to demonstrate the ability to incinerate oil and hazardous materials on-site. Following several years of design and construction, a unit was tested in 1981.

The MIS consists of a 16-foot rotary kiln with a dual refractory lining. This is connected to an 18-foot by 50-inch-diameter cyclone, which feeds into a secondary chamber (42 feet long by 52 inches i.d.) and quench sump. The exhaust then enters a wet electrostatic precipitator which is 40 feet high with a cross section that is 84 inches by 84 inches. The final unit is a mass transfer scrubber 8 feet long by 5 feet high.

The entire system can be mounted on over-the-road equipment and transported to a cleanup site. A total of 21 trailers are required to transport this incineration system.

Types of Wastes Handled. During trial burns in 1982 and 1983 the MIS successfully incinerated PCBs and other chlorinated organic liquids. The unit was subsequently modified to handle solid materials and field-tested at Denny Farm in southwestern Missouri. In this test the incinerator was fed

with soil contaminated with 2,3,7,8-tetrachlorodibenzo(*p*)dioxin (2,3,7,8-TCDD) and related chemicals. This test was successfully completed in April 1985. Between 1985 and 1989 the incinerator treated over 12,500,000 pounds of solids and 230,000 pounds of liquids (EPA, Sept. 1990).

Phase Separation

Phase separation is used for separating solid/liquid or liquid/liquid suspensions with different specific gravities. It encompasses many different processes, several of which are described below.

Oil Separation. This process employs a number of equipment configurations to remove oil from water, each of which provides surface contact for de-emulsifying the oil particles from the water phase. A coalescer is a flow-through chamber containing metal sheets inclined at an angle, causing the oil droplets to agglomerate and form a continuous oil phase. The lighter oil fraction rises to the top of the chamber and is skimmed off.

Centrifugation. In this process the components of the oil/liquid mixture are separated mechanically by applying centrifugal force. This is done by rapidly rotating the mixture in a confined vessel. The suspended oils will migrate toward the center of the vessel. Centrifuges also may be used for separation of liquid/solid mixtures.

Dissolved Air Flotation (DAF). These processes are used to separate emulsified oils from water by first dissolving air (under high pressure) into the water and then dropping the pressure. This causes tiny air bubbles to be generated throughout the water phase. The oils accumulate at the air–water interface and are carried to the top of the chamber where they are skimmed off. DAF units usually are employed for more complete oil removal, whereas coalescers are used for coarse oil separation.

Operating Parameters. Mobile phase separation units depend on the following parameters:

- Oil concentration
- Detention time
- Surface area
- Skimming rate
- Air pressure
- Revolutions per minute (rpm)
- Treatment chemical needed

Waste Type Handled. The following wastes can be treated with phase separation:

- Immiscible oily liquids in water
- Suspended solids
- Hydrophobic chemicals
- Two-phase leachates

Ozonation

Ozone is an unstable gas that is slightly soluble in water (having about 13 times the solubility of oxygen). Although relatively stable in air, in aqueous solution ozone continues to decompose at increasing rates with increasing concentrations of oxidizable contaminants. Thus, solutions of ozone do not provide long-lasting residuals.

Ozone is the most powerful oxidizing agent/disinfectant readily available for water and wastewater treatment. As a result, it is used for the following purposes:

- Disinfection and viral inactivation
- Chemical oxidation
- Preoxidation in preparation for biological treatment

The amounts of ozone needed to perform each of these functions depend upon a number of factors, but primarily upon the ozone demand of the constituents of the water/wastewater to be ozonized.

During preozonation to prepare solutions or during subsequent biological treatment, only small amounts of ozone are added (1 mg/l of ozone per mg/l of dissolved organic carbon). When it is used to oxidize pollutants that react rapidly with ozone, attainment of

measurable ozone concentrations can be taken as achievement of the ozone demand.

A concentration of residual ozone can be achieved and monitored during disinfection and/or viral inactivation of waters that have little extraneous ozone demand (drinking water, swimming pool water, cooling water, etc.). For other drinking water or industrial water treatment applications, and for most wastewater applications, control of ozonation processes must be monitored by a surrogate analytical technique. Such process controls are not based upon the monitoring of dissolved residual ozone.

Lagoons and Septic Tanks

Many well-populated areas in the United States are not served by sewer systems or waste treatment plants. Lagoons and septic tanks are the usual alternatives in such situations.

Lagoons, or, as they are sometimes called, stabilization or oxidation ponds, have several advantages when used correctly. They can provide primary and secondary treatment of sewage, or they can be used to supplement other processes. A lagoon is a specially designed pond, usually 3 to 5 feet deep, in which sunlight, algae, and oxygen interact to restore water to a quality equal to or better than that of the effluent from secondary treatment. Although changes in the weather will affect how well lagoons break down sewage, they can be very effective when used with other waste treatment processes.

Solvent Extraction Process

The Basic Extraction Sludge Treatment (B.E.S.T.®) process developed by the Resources Conservation Company is a solvent extraction process used to treat contaminated soils and sludges. This patented process uses an aliphatic amine to break down suspensions and emulsions and separate a waste into three constituents—dischargeable water, oil and organics, and dry, oil-free solids. Treatment costs vary with feed characteristics but generally have been demonstrated to be lower

Fig. 3.8. B.E.S.T.® soil cleanup unit schematic. (*Courtesy Resources Conservation Company, Inc., Bellevue, WA.*)

than those for incineration or commercial landfilling. (See Fig. 3.8.)

The process can be used to treat wastes from refineries, petrochemical plants, wood treating operations, wastewater treatment systems, and steel mills. In addition, PCB-contaminated sludges and soils and the RCRA-listed hazardous wastes (shown below) are suitable feed stocks.

K001*—Creosote sludges
K048—OAF float
K049—Slop oils emulsion solids
K050—Heat exchanger sludge
K051—API separator bottoms
K052—Leaded tank bottoms
K062—Ammonia still sludge
K089—Decanter tar sludge
K101—Distillation residues
K044—Water treatment sludges

KPEG Chemical Dechlorination Process

Chemical dechlorination processes that use metallic sodium to treat materials contaminated with polychlorinated biphenyls (PCBs) are considered proven technologies. However, metallic sodium is difficult to handle and dangerous. Dr. L. Pytlowski at the Franklin Institute Research Laboratory is credited with initially discovering that alkaline-glycolate reagents could detoxify chlorinated aroniaties, especially polychlorinated biphenyls.

It is generally accepted that the initial reaction (between, e.g., potassium polyethylene glycolate and PCB) is a nucleophilic substitution that produces an aromatic, as shown in the following equation:

$$ROH + KOH \longrightarrow ROK + HOH$$

This reaction can be conducted at moderate temperatures and ambient pressure in conventional equipment. Water will slow but not stop the reaction.

Soil decontamination tests have demonstrated that the process can effectively decontaminate soils containing up to 10,000 ppm of PCBs and soils or sludge containing 5 percent or less of organics; soils with higher contaminant concentrations may be more economically treated by incineration. However, this process can treat soils containing more than 5 percent organics. (See Fig. 3.9.)

Bioremediation of Wastewater

Various biological materials are found in contaminated wastewaters. The removal mechanisms include adsorption, bioaccumulation, metal reduction, and conversion to insoluble metal sulfides. Normally, most microbial techniques are specific for individual metals.

Researchers have been successful with a microbial strain, *Bacillus megaterium*, that has the ability to adsorb a variety of heavy metals. In its dormant spore form, it has an affinity for many heavy metals. Additionally, this form of the organism does not require nutrients and is resistant to the usual toxic effects of heavy metals.

Tests with *Bacillus megaterium* isolated from sedimentation contaminated with mercury, lead, strontium, and cesium have been conducted on laboratory solutions and electroplating rinse waters. The rinse water pH was adjusted with HCl, and the solution was filtered. Spores were added to give a 0.10 weight ratio (wet) of spores to rinse water, and the mixture was incubated for 24 hours.

*EPA Office of Solid Waste Code Classification for the named waste type.

Fig. 3.9. KPEG decontamination process for PCB soils.

The spores then were filtered from the mixture, and HCl was added to adsorb the metals from the spores.

Some competition for the bonding sites between metal species seeking the same type of bonding site have been noted. For example, silver, arsenic, and cadmium appeared to seek similar adsorption sites. When one of these species was added to a solution containing another, adsorption of the latter was correspondingly decreased.

Researchers at Worcester Polytechnic Institute are investigating the feasibility of developing a wastewater treatment technology using a fungus called *P. ochro-chloron.* This fungus, discovered growing in an electroplating wastewater stream in Japan, is obtained from spores grown in petri dishes. The organism produces a cell wall that has the ability to bind large amounts of heavy metals as divalent cations. Heavy metal ions are added to the growth medium to ensure that the fungus grows in colonies that form beadlike structures.

After the fungi beads are grown, they are killed by using ethanol and then treated with a mixture of sodium carbonate and bicarbonate. This treatment prepares the fungal cell walls to exchange sodium ions for heavy metal ions. The fungal beads are packed into a column, and metal-containing water is passed through the column. The adsorbed heavy metals can be retrieved from the biological material by washing the beads with a dilute acid solution. This work is in the early stages of commercialization.[1]

Biological treatment processes have been gaining increasing respect in the hazardous waste industry as viable treatment alternatives. Although the number of these processes commercially available for hazardous waste treatment is still small, a number of new biological treatment processes are on the commercial threshold. One such process utilizes a common white rot fungus, *Phanerochaete chrysosporium,* which has the ability to degrade trichlorophenol, trinitrotoluene (TNT), DDT, lindane, PCBs, benzo[α]pyrene, dioxins, and other chlorinated organics.

Investigation of various biological reactor configurations revealed that a fixed-film reactor is compatible with the fungus. A rotating biological contactor was selected because this design is already in widespread commercial use and thus requires no further development.

The rotating biological contactor is first charged with a concentrated *P. chrysosporium* spore suspension. Typically, the suspension comprises 20 percent of the reactor volume although it has been found that the amount can be reduced by a factor of 100 with no adverse effects. The fungus is allowed to grow in the reactor for 2 to 4 days until the reactor surfaces are completely covered. During this period, nutrient solution is added to promote growth.

At this point, the reactor conditions are altered to force the fungus into a secondary metabolic state. In this state, the fungus excretes enzymes capable of degrading a wide variety of organic compounds by catalyzing

oxidation reactions. Because the fungus does not excrete these enzymes unless it is nutritionally stressed, one or more of a variety of nutrients, usually nitrogen, are withheld from the medium to induce the secondary metabolic state.

The optimum temperature for the process appears to be 102°F (39°C), and the pH is typically 4.5. The oxygen concentration should be at least 20 percent in the enclosed space above the liquid in the reactor, and 50 to 100 percent oxygen is preferable.

Once the secondary metabolic state has been induced, the reactor is ready to accept aqueous waste for treatment. The process can be operated in either a batch or a continuous mode. Effluent from the reactor must undergo additional separation and filtration steps prior to reuse or discharge.

Organic compounds are eventually degraded to water, carbon dioxide, and inorganic ions by the process. No significant sludge disposal problems are associated with the process because very little fungal growth occurs after the initial growth period.

In bench-scale tests, pure TNT was degraded from an initial concentration of 90 ppm to 5 ppm in 8 hours. TNT concentrations in both industrial and synthetic "pink water" (i.e., wastewater from the manufacture of TNT munitions) were reduced from 90 ppm to nondetectable levels in less than one day. Other tests indicate that 50 ppm of 2,4,6-trichlorophenol also can be degraded in less than one day.

TYPES OF INDUSTRIAL WASTES

Table 3.2 presents the Standard Industrial Classification (SIC) of industries of significance for water pollution. The great variety of industrial pollutants argues against an attempt to catalog them. However, for the purpose of general description, at least five distinct categories of waste from industrial sources can be recognized:

1. Oxygen-demanding materials.
2. Settleable and suspended solids.
3. Many materials that impart acidity or alkalinity.
4. Heat.
5. Toxic compounds.

Within these general categories of pollutants are included most of the possible sources of recognized water-quality problems. Only bacterial and viral presences fall outside the group of pollutants to which industry contributes materially. (However, even they are not entirely outside the scope of the polluting activities of industry. For example, meat packing plants, and other food processors in lesser measure, contribute to the presence and viability of water-borne bacteria.) Manufacturing must stand near the top of, and very probably leads, any list of potential sources of water pollution. (See Table 3.3.)

Therefore, in any possible strategy of water-pollution control, industrial wastes are of critical importance. Over the last few years, the amount and composition of industrial output has been such that for every new increment of BOD that has been generated directly by population increase, 20 more units have been generated by increased industrial output. Increased per capita production is the essence of improvement in a society's standard of living, and the production of wastes is an inescapable concomitant of the production of goods. So as population and per capita production continue to advance, we can anticipate a continuing and unavoidable advance in the volume of wastes that must be managed.

Fortunately, industry has added rapidly to its inventory of waste-treatment facilities since the mid-1970s, and it appears that provision for waste treatment is routinely designed into new plants and plant additions. Even though additional treatment of municipal wastes has had an undesirable influence on the nation's ability to contain the level of waste discharges, the preponderence of industrial wastes and their extremely rapid rate of increase would have made containment impossible if increases in industrial-waste-treatment effectiveness had not outdistanced production increases.

TABLE 3.2 Standard Industrial Classification of Industries of Significance for Water Pollution

CODE	
20	FOOD AND KINDRED PRODUCTS
201	Meat products
2011	Meat slaughtering plants
2013	Meat processing plants
2015	Poultry dressing plants
202	Dairy products
2021	Creamery butter
2022	Natural and process cheese
2023	Condensed and evaporated milk
2026	Fluid milk
203	Canned and frozen foods
2033	Canned fruits and vegetables
2034	Dehydrated food products
2035	Pickled foods, sauces, salad dressings
2037	Frozen fruits and vegetables
204	Grain mill products
2041	Flour and other grain mill products
2043	Cereal preparations
2046	Wet corn milling
205	Bakery products
206	Sugar
207	Candy and related products
208	Beverage industries
2082	Malt liquors
2084	Wines and brandy
2085	Distilled liquors
2086	Soft drinks
209	Miscellaneous foods and kindred products
2091	Cottonseed oil mills
2092	Soybean oil mills
2094	Animal and marine fats and oils
2096	Shortening and cooking oils
22	TEXTILE MILL PRODUCTS
2211	Weaving mills, cotton
2221	Weaving mills, synthetic
2231	Weaving, finishing mills, wool
225	Knitting mills
226	Textile finishing, except wool
228	Yarn and thread mills
229	Miscellaneous textile goods
24	LUMBER AND WOOD PRODUCTS
2421	Sawmills and planning mills
2432	Veneer and plywood plants
2491	Wood preserving

CODE	
26	PAPER AND ALLIED PRODUCTS
2611	Pulp mills
2621	Paper mills, except building
2631	Paperboard mills
264	Paper and paperboard products
265	Paperboard containers and boxes
2661	Building paper and building board mills
28	CHEMICALS AND ALLIED PRODUCTS
281	Basic chemicals
2812	Alkalies and chlorine
2818	Organic chemicals, n.e.c.
2819	Inorganic chemicals, n.e.c.
282	Fibers, plastics, and rubbers
2821	Plastics materials and resins
2823	Cellulosic man-made fibers
2824	Organic fibers, noncellulosic
283	Drugs
284	Cleaning and toilet goods
2851	Paints and allied products
2861	Gum and wood chemicals
287	Agricultural chemicals
289	Miscellaneous chemical products
29	PETROLEUM AND COAL PRODUCTS
2911	Petroleum refining
295	Paving and roofing materials
30	RUBBER AND PLASTICS PRODUCTS, n.e.c.
3069	Rubber products, n.e.c.
3079	Plastics products, n.e.c.
31	LEATHER AND LEATHER PRODUCTS
3111	Leather tanning and finishing
32	STONE, CLAY, AND GLASS PRODUCTS
3211	Flat glass
3241	Cement, hydraulic
325	Structural clay products
326	Pottery and related products
327	Concrete and plaster products
3281	Cut stone and stone products
329	Nonmetallic mineral products
33	PRIMARY METAL INDUSTRIES
331	Steel rolling and finishing mills
332	Iron and steel foundries
333	Primary nonferrous metal
3341	Secondary nonferrous metals

The processes that can be used to separate heavy metals from industrial liquid waste streams are given in Table 3.4, processes that can treat various tars or solids are shown in Table 3.5, and processes suitable for treating slurries and sludges are listed in Table 3.6.

The EPA national screening survey of hazardous waste treatment, storage, disposal, and recycling facilities yielded the data shown in Table 3.7. The compounds listed in Table 3.8 are refractory industrial waste compounds.

A summary of potentially suitable industrial wastewater treatment technologies for various waste types is given in Table 3.9. Additionally, there are three basic approaches to land treatment of nonhazardous industrial effluent wastewater, as depicted in Fig. 3.10.

INDUSTRIAL WASTE MANAGEMENT STRATEGIES

Unlike the public sector, where only variations on a single theme of waste treatment are possible, industrial pollutants can be managed through at least four distinct strategies:

1. Installation of waste treatment plants, the most obvious procedure.
2. Discharge of industrial wastes to public sewage systems for treatment. (Industry is estimated to account for half of the current BOD loading to metropolitan area waste treatment plants.)
3. Process modification.
4. Changed product formulations.

Strategies (3) and (4) together comprise the most effective as well as the most efficient means of reducing wastes.

Joint Treatment

Joint treatment is an alternative that can be advantageous both to a publicly owned treatment works (POTW) and to industry. Treatment of industrial wastewaters is incidental to a POTW's primary function of treating domestic sewage. Where the industrial contribution constitutes a significant portion of the total flow and substantially alters the concentration of pollutants normally contained in domestic sewage, the public agency may resort to a joint treatment approach. Then the industry or industries contributing the pollutants are made partners in the design and construction of the system, and the treatment works is designed specifically to remove the industrial pollutants. Both capital costs and operating costs are allocated to the industry and the public agency according to an agreement arrived at through negotiation, or as required by federal regulations if construction grant funds are involved.

Joint treatment of industrial wastewaters with municipal domestic sewage offers the following advantages:

- Increased flow, which can result in reduced ratios of peak to average flows.
- More efficient use of land resources, particularly in cases where available land for treatment facilities is scarce.
- Improved operations (larger plants are potentially better operated than smaller plants).
- Increased number of treatment modules with resultant gains in reliability and flexibility.
- More efficient disposal of sludges resulting from treatment of wastewaters containing pollutants susceptible to treatment in POTWs.
- Utilization of the nutrients available in domestic wastes for biological treatment of industrial wastes that are nutrient-deficient.

There also may be disadvantages to joint treatment:

- Where the pollutants are different from those usually treated in a POTW, a design to treat the combined industrial/domestic waste stream for those pollutants may not be cost-effective.
- Joint treatment by definition implies that the POTW was designed so as not to be interfered with by industrial wastes. However, where it requires design modifications ordinarily not required for domestic

TABLE 3.3 Wastewater Characteristics and Pollutants of Selected Industry

Liquid Waste Characteristic	*Domestic*	*Meat Products*	*Canned and Frozen Foods*	*Sugar*	*Textile Mill Products*	*Paper and Allied Products*
Unit volume	x	x	x	x	x	x
pH		x		x	x	x
Acidity				x	x	
Akalinity		x		x	x	x
Color					x	x
Odor					x	
Total solids		x		x	x	x
Suspended solids	x	x	x	x	x	x
Temperature		x				x
BOD_5/BOD ultimate	x	x	x	x	x	x
COD	x					
Oil and grease		x		x	x	x
Detergents (Surfacants)	x				x	x
Chloride		x				x
Heavy metals						
Cadmium						
Chromium					x	
Copper					x	
Iron						
Lead						
Manganese						
Nickel						
Zinc					x	
Nitrogen						
Ammonia		x		x	x	x
Nitrate					x	
Nitrite					x	
Organic		x		x		x
Total		x			x	x
Phosphorus		x				
Phenols					x	
Sulfide					x	x
Turbidity	x				x	
Sulfate					x	x
Thiosulfate					x	
Mercaptans						x
Lignins						x
Sulfur						x
Phosphates						x
Potassium						x
Calcium						x
Polysaccharides						x
Tannin						x
Sodium						x
Fluorides						
Silica						
Toxicity					x	
Magnesium						
Ammonia						
Cyanide						
Thiocyanate						
Ferrous iron						
Sulfite						x
Aluminum						
Mercury						x

[a]Source: *The Total Cost of Clean Water*, Volume II, FWPCA, U.S. Department of the Interior,

Basic Chemicals	*Fibers Plastics and Rubbers*	*Fertilizer*	*Petroleum Refining*	*Leather Tanning and Finishing*	*Steel Rolling and Finishing*	*Primary Aluminum*	*Motor Vehicles and Parts*
x	x	x	x	x	x	x	x
x	x	x	x	x	x	x	x
x		x	x	x	x	x	x
x			x	x	x		x
x	x		x				x
x	x		x				x
x	x	x		x	x	x	x
x	x		x	x	x		x
x	x	x			x		x
x	x		x	x	x		x
x	x		x		x		x
x	x		x		x		x
x	x						x
x	x		x	x	x		x
x							x
x							x
x							x
x			x		x		x
x							x
x			x				x
x							x
x							x
x		x	x	x	x		
x							
x							
x			x	x			
x		x		x			
x		x					
x	x		x		x		x
x			x	x	x		
x			x				
x	x		x	x			
x							
x			x				
x							
x		x					x
x			x				
x			x				
x							
				x			
x			x	x		x	
x		x	x		x	x	
x		x					x
x	x						
			x				
x		x	x		x		
x				x	x		x
x					x		
x					x		
x							
x							
x		x	x				

U.S. Government Printing Office, Washington, DC. Apr. 1, 1968.

TABLE 3.4 Processes That Separate Heavy Metals from Liquid Waste Streams

Process	*Required Feed Stream Properties*	*Characteristics of Output Stream(s)*
Physical Removal		
Ion exchange	Con. <4000 ppm; aqueous solutions, low SS	One concentrated in heavy metals; one purified
Reverse osmosis	Con. >400 ppm; aqueous solutions; controlled pH; low SS; no strong oxidants	One concentrated in heavy metals; one with heavy metal concentrations >5 ppm
Electrodialysis	Aqueous solutions; neutral or slightly acidic; Fe and Mn <0.3 ppm; Cu <400 ppm	One with 1000–5000 ppm heavy metals; one with 100–500 ppm heavy metals
Liquid ion exchange	Aqueous solutions; no concentrations limits; no surfactants; SS <0.1%	Extraction solvent concentrated in heavy metals; purified water or slurry
Freeze crystallization	Aqueous solutions; TDS <10%	Concentrated brine or sludge; purified water, TDS ~100 ppm
Chemical Removal		
Precipitation	Aqueous or low viscosity non-aqueous solutions, no concentration limits	Precipitated heavy metals sulfides, hydroxides, oxides, etc.; solvent With TDS governed by solubility product of precipitates
Reduction	Aqueous solutions; concentrations of heavy metals <1%; controlled pH	Acidic solutions with reagant (Oxidized Na_2PO_4 or Zn); metallic precipitates
Electrolysis	Aqueous solutions; heavy metal concentrations <10%	Recovered metals; solution with 2–10 ppm heavy metal

*Source: *Pollution Engineering*, Feb. 1988.

TABLE 3.5 Processes That Can Treat Tars or Solids

Process	*Required Feed Stream Properties*	*Characteristics of Output Stream(s)*
Calcination	Tars or solids that can be volatilized	Volatile; char and/or metal oxides
Hydrolysis	Tars or solid powders	Hydrolysis products
Steam distillation	Solids contaminated with volatile organics	Purified solids; condensed organics
Dissolution	Tars, solids, or solid powders that will dissolve in some reagant	Liquid solution for further treatment; solid residue
Crushing and grinding	Bulk solid	Powdered solid
Cryogenics	Tar, bulk solid	Reduced particle size

*Source: *Pollution Engineering*, Feb. 1988.

wastes, joint treatment may not be cost-effective.

- If joint treatment results in sludge disposal or utilization problems, it may not be acceptable.

About 16,000 municipal wastewater treatment facilities treating 37 B gallons per day, many commercial establishments, and 160,000 industries discharge their wastes into POTWs; about 39,000 industries discharge directly into waterways.

Wastewater treatment is primarily a local function, with 15,438 systems (*EPA 1988 Needs Survey* data). (See Table 3.10.)

The escalating cost of upgrading, maintaining, and operating the existing wastewater treatment infrastructure, expanding it to support residential and commercial growth, and providing a new infrastructure where it

TABLE 3.6 Processes That Can Treat Slurries or Sludges

Process	*Required Feed Stream Properties*	*Characteristics of Output Stream(s)*
Calcination	Waste stream with components that decompose by volatilization (hydroxides, carbonates, nitrates, sulfites, sulfates)	Solid greatly reduced in volume, volatiles
Freeze crystallization	Low-viscosity aqueous slurry or sludge	Brine sludge; purified water
HGMS	Magnetic or paramagnetic parties in slurry	Particles adsorbed on magnetic filter
Liquid ion exchange	Solvent-extractable inorganic component	Solution in extraction solvent
Flotation	Floatable particles in slurry	Froth
Hydrolysis	Hydrolyzable component	Hydrolysis products
Anaerobic digestion	Aqueous slurry <7% solids; no oils or greases; no aromatics or long-chain hydrocarbons	Sludges, methane, and CO_2
Composting	Aqueous sludge; <50% solids	Sludge; leachate
Steam distillation	Sludge or slurry with volatile organics	Volatile, solid residue
Solvent extraction	Solvent-extractable organic	Solution of extracted components; residual sludge

Source: *Pollution Engineering*, Feb. 1988.

TABLE 3.7 Hazardous Waste Managed by Industry

Industry	*Number of Facilities*	*Amount Waste Managed (MMT)**	*General Waste Description*
Chemical	700	218	Contaminated wastewaters, spent solvent residuals, still bottoms, spent catalysts, treatment sludges, and filter cakes
Fabricated metals	200	4	Electroplating wastes, sludges contaminated with metals and cyanides, degreasing solvents
Electrical equipment	240	1	Degreasing solvents
Petroleum refinery	100	20	Leaded tank bottoms, slop oil emulsion solids, API separator sludge, DAF float
Primary metals	150	4	Pickle liquor, sludge with metal contaminates
Transportation equipment	150	3	Degreasing solvents, metals, sludges
National security	100	1	All types of wastes
Other	1360	24	All types of wastes
Total	3000	275	

*Million metric tons.
Source: National Screening Survey of Hazardous Waste Treatment, Storage, Disposal and recycling Facilities, U.S. EPA, Office of Solid Waste, Office of Policy, Planning and Information (1986).

does not exist is creating problems in the United States. Additional requirements for the treatment of toxics, in addition to the conventional pollutants further complicate the issue. Toxic chemical control also dictates that the treatment, handling, and disposal of treatment process sludge or residues receive considerably more attention than they have had in the past.

Sludge Management. Sludge management represents about half of wastewater treatment operation and maintenance (O&M) costs. The ultimate disposal of sludges produced by either pretreatment or joint treatment operations is an important consideration. The POTW must be aware of effects on the environment that may result from sludge disposal. Pretreatment facilities normally

TABLE 3.8 Refractory Industrial Wastes

Nitrobenzene[a]	Bromobenzene
Trichloroethane	Dichlorobenzene
Tetrachloroethylene	Bromochlorobenzene
Chloroethyl ether	Ethylbenzene[a]
Chloromethyl ethyl ether	Chloroform
Chloropyridine	Styrene[a]
Chloronitrobenzene	Isopropylbbenzene
Dichloroethyl ether	Butylbenzene
Benzene[a]	Dibromobenzene
Toluene[a]	Isocyanic acid[a]
Camphor	Methylchloride[a]
Veratrole (1,2-dimethoxy benzene)	Bromophenylphenyl ether
Guaiacol (methoxy phenol)	Dinitrotoluene[a]
Borneol (bornyl alcohol)	Methylbiphenyl
Isoborneol	Acetone
Ethylene dicloride[a]	2-Ethylhexanol
Chlorobenzene	2-Benzothiozole

[a]These compounds also have been found to impart taste and odor to drinking water supplies in trace amounts.

remove incompatible pollutants that may be deposited in the POTW sludge, and their removal can be an advantage in terms of the environmental problems connected with ultimate sludge disposal. Incompatible pollutants in sludges can cause problems with most disposal techniques including incineration, landfills, ocean dumping, and land spreading. Consequently, the removal of incompatible pollutants at their source by pretreatment is advantageous.

However, incompatible pollutants removed by pretreatment still require ultimate disposition. The sludge produced by pretreatment operations may be a source for by-product recovery or recycle; but when this is not economically or technically feasible, disposal of the sludge is necessary. Although the sludges produced by industrial pretreatment may not technically be under municipal regulatory control, their impact on other environmental areas should be noted.

Inhibition of Biological Treatment Processes. Interfering substances are generally categorized as inorganics and organics. Acidity, alkalinity, pH, ammonia, metal ions, sulfate, and sulfide comprise the major inorganic interferences. Organic substances include alcohols, amines, ketones, aldehydes, chlorinated (or halogenated) hydrocarbons, pesticides, herbicides, phenols, surfactants, and miscellaneous (both aromatic and aliphatic) chemicals.

A precipitated pollutant has relatively little inhibitory impact on biological treatment compared to soluble compounds. Settleable solids usually are removed in the primary treatment phase and therefore seldom reach the secondary or biological treatment phase.

Whether or not a substance is inhibitory depends on factors such as concentration and the presence of other chemicals that have synergistic or antagonistic effects. Some substances, such as mercury, even when present in wastewater at a very low concentration can disrupt one or more functions of a biological treatment system. Others, such as chloride ion, are inhibitory only at relatively high concentrations. Special phenomena that may alter the inhibitory nature of a substance include synergism, antagonism, and acclimation. These are described below.

Synergism is characterized as an increase in the inhibiting effect of one substance by the presence of another. Synergism and its opposite, antagonism, are most prevalent in

TABLE 3.9 Potentially Suitable Industrial Wastewater Treatment Technologies

Waste Type	*Incineration*	*Pyrolysis*	*Wet Oxidation*	*Neutralization*	*Precipitation*	*Distillation*	*Air Stripping/ Soil Aeration*	*Activated Carbon*	*Evaporation/ Dewatering*	*Phase Separation*	*Fixation*	*Extraction/Soil Washing*	*Membrane Sep./ Ion Exchange*	*Evaporation*	*Filtration*	*Activated Sludge*	*In situ Biodegradation*
Aqueous Wastes:																	
Metals	×	×	×	●	●	×	×	○	×	×	×	○	●	●	●	×	×
Highly toxic organics	○	○	●	○	○	●	×	●	×	●	×	○	●	●	●	○	○
Volatile organics	○	○	●	○	○	●	●	●	×	●	×	○	○	×	×	○	○
Toxic organics	○	○	●	○	○	●	●	●	×	●	×	○	●	●	×	●	●
Radioactive	×	×	×	×	○	×	×	×	×	×	×	×	×	○	●	×	×
Corrosive	×	×	×	●	●	○	×	×	×	×	×	×	×	×	×	×	×
Cyanide	○	○	●	○	○	×	×	×	×	×	×	×	×	×	×	○	×
Pesticide	○	○	●	●	●	●	×	●	×	●	×	○	●	●	●	○	×
Asbestos	×	×	×	○	○	×	×	×	×	×	×	×	○	●	●	×	×
Explosive	○	×	×	×	●	×	×	●	○	●	×	○	●	○	×	○	○
Organic Liquids:																	
Metals	○	○	×	●	●	●	×	×	×	●	○	×	×	●	●	×	×
Highly toxic organics	●	●	●	×	×	●	×	○	×	○	×	×	×	×	×	×	×
Volatile organics	●	●	●	×	×	●	×	×	×	○	×	×	×	×	×	×	×
Toxic organics	●	●	●	×	×	●	×	×	×	○	×	×	×	×	×	×	×
Radioactive	○	×	×	×	○	●	×	×	×	○	○	×	×	○	○	×	×
Corrosive	×	×	×	●	○	●	×	×	×	○	×	×	×	×	×	×	×
Cyanide	●	●	●	×	×	●	×	×	×	○	×	×	×	×	×	×	×
Pesticide	●	●	●	×	×	●	×	×	×	○	×	×	×	×	×	×	×
Sludges/Soils:																	
Metals	×	○	×	○	●	×	×	○	●	×	●	●	×	○	×	×	×
Highly toxic organics	●	●	●	×	×	●	×	●	●	●	●	○	×	●	●	×	○
Volatile organics	●	●	●	×	×	●	●	○	●	○	○	○	×	×	×	×	○
Toxic organics	●	●	●	×	○	○	×	●	●	●	●	○	×	○	○	×	●
Radioactive	×	×	×	×	×	×	×	×	●	●	●	×	×	×	×	×	×
Corrosive	×	×	×	●	×	×	×	×	●	×	●	×	×	×	×	×	×
Cyanide	○	○	●	×	○	○	×	×	●	×	○	×	×	×	×	×	○
Pesticide	●	●	○	×	●	○	×	●	●	○	●	●	×	●	○	×	○
Asbestos	×	×	×	×	×	×	×	×	●	×	●	×	×	×	×	×	×
Explosive	○	○	×	○	×	×	×	●	●	●	●	○	×	×	×	×	○

● Applicable ○ Potentially applicable × Not applicable

Source: EPA Office of Solid Waste and Emergency Response, 540/2-86/003(f), Sept. 1986. (*Courtesy EPA.*)

situations involving combinations of transition metals or heavy metals. The inhibitory effects also are enhanced by acidity.

The synergistic effects of such metals in combination with acidity can be understood from the chemistry of these metals. In the usual pH range of sewage influents (somewhat basic), heavy metals tend to be insolubilized by hydrolysis. They precipitate or adsorb on solids, and they interact with polyelectrolytes or various chemical species containing anionic functional groups. Acidity suppresses hydrolysis, keeping the metal ions in solution, and the hydrogen ions compete with the metal ions for adsorption on solids or anionic functional groups in solution.

The synergism of cyanide or other complexing substances makes metal wastes more easily biodegradable. It is possible for a microorganism to ingest excessive levels of

Fig. 3.10. Three basic approaches to treatment of nonhazardous waste waters. (*Courtesy EPA.*)

TABLE 3.10 Distribution Information; Facilities and Treatment

Size (MGD)	*No. of Municipal Plants*	*Capacity in 10^3 gallons*	*Percent*
0.01– 0.1	4960	251	0.7
0.11– 1.0	7003	2671	7.2
1.01–10.0	2893	9372	25.6
10.01+	577	24,383	66.5
Totals	15,438	36,677	100.0
Treatment			
Less than 2nd	2122	5521	15.1
Secondary	8403	15,714	42.6
More than 2nd	3115	14,373	39.2
No discharge	1762	973	2.7
Unknown	46	88	0.2
Totals	15,438	36,677	100.0

*According to *EPA 1988 Needs Survey* data, wastewater treatment is primarily a local function with 15,438 systems.

complexed metal ion and then to be destroyed by assimilating the complexing substance that shields the microorganism from the metal ion. The subsequent release of an excessive level of the metal within the organism upsets its biological life processes.

Antagonism is the opposite of synergism; it is characterized as a decrease in the inhibitory effect of one substance caused by the presence of another. Antagonistic effects occur with combinations of metallic and anionic pollutants. Chelating agents such as EDTA (disodium salt of ethylenediamine tetraacetic acid) and HEDTA (disodium salt of hydroxyethylenediamine triacetic acid) exhibit antagonistic properties with metals. Chelating agents are used in culturing microorganisms to regulate the level of metals needed to grow bacterial cultures. Bacteria survive in culture solutions containing concentrations of these metals high in the inhibitory range when chelating agents are present.

There is ambiguity in the literature about the inhibitory effects of metals on sludge digestion. This ambiguity may be explained by the antagonistic effect of the sulfide normally present in a digester. Sulfide ion precipitates metals, removing them from solution and thus eliminating their inhibitory effect. Knowledge of this behavior has been used effectively in putting upset digesters back into operation. Additionally, other ions such as hydroxide, chromate, ferricyanide, phosphate, carbonate, and arsenate will tend to precipitate with metals, thereby reducing their inhibitory effects.

One must consider significant biological factors to understand the inhibitory effect of pollutants. In the activated-sludge process, a healthy biomass contains a broad distribution of microorganisms, including many species of bacteria and protozoa. In the absence of sufficient food, reproduction slows, and the microorganisms devour each other. Whenever the environment changes because of the introduction or the omission of a given pollutant, the opportunities for reproduction and growth of different species change.

The history of an activated-sludge biomass affects the way in which it will respond when a new pollutant is introduced; species that cannot tolerate the substance fail to reproduce and grow and tend to die off, whereas more tolerant species consume the food supply and grow and reproduce.

When a biomass becomes accustomed to the presence of a normally inhibitory concentration of a substance, it can be characterized as *acclimated* to that pollutant. But sludge digestion and nitrification do not have the

same flexibility of adaptation to changing environmental conditions that other biological processes have. Both nitrification and sludge digestion are biological processes that rely on particular strains of bacteria. Sludge digestion proceeds in two steps, using two specific bacterial strains to achieve digestion. Nitrification is also limited to particular bacterial types. When adverse conditions are encountered in these processes, there is no possibility of another organism taking over for the affected strain of bacteria. Neither nitrification nor sludge digestion is readily acclimated to a new pollutant, and both may be easily upset.

A POTW may contain many thousands of inorganic compounds present as major or minor constituents. Most inorganic substances dissolved in wastewater are present in ionic form, and the number of parameters of interest can be reduced to a few cations and anions and nonionics.

The major cations consist of the ammonium ion and various metal ions. A few metal ions, chiefly sodium, calcium, and magnesium, are prevalent in wastewaters but are not inhibitory except at very high concentrations

SUPERFUND INNOVATIVE TECHNOLOGY EVALUATION (SITE) PROGRAM

The EPA initiated this program to promote and evaluate newly developed techniques that had not been fully demonstrated. The goal of this effort was to find new ways to remediate uncontrolled hazardous waste sites. The technologies accepted for field-scale evaluation represented a wide array of innovative methods, from thermal treatment and bioremediation to soil washing, solvent extraction, and in situ stripping.

The purpose of the SITE Program Applications Analysis Reports is to provide overall assessment of the performance, reliability, and costs of technologies in the Demonstration Program, relative to typical Superfund sites. Reports utilize all available data including the results of the SITE demonstration, the vendor's design and test data, and other laboratory and field applications of the technology. An analysis is given of the advantages, disadvantages, and limitations of the technology. Cost estimates are based on pilot-scale and commercial applications and available data. The report also examines factors, such as site and waste characteristics, that have a major impact on costs and performance. Highlights of Applications Analysis for some of the technologies demonstrated are provided below.

In the solidification/stabilization process developed by Hazcon, Inc., a patented binding agent is mixed with pozzolans in a mobile, truck-mounted blending unit.[2] The resulting concrete mass has been shown to enhance solidification of wastes containing high concentrations of organics and to immobilize heavy metals. TCLP leachate reduction for metals often is achieved in excess of 100 ppm. However, organic contaminants (volatiles and semivolatiles) were not immobilized based on SITE TCLP tests. Tests indicated that the permeabilities of the treated soil were very low, durability tests were good, and no loss of strength was observed after a weathering test. The resulting volume of the treated soil was approximately double that of the undisturbed waste feed. Although the equipment experienced some operating difficulties during the field demonstration, it is expected to operate efficiently with some design changes. The technology is economical, with costs estimated to be between $97 and $207 per ton.[3]

The Electric Infrared Incineration process, originally developed by Shirco Infrared Systems, Inc., is an electric infrared thermal processing system that uses silicone-carbide rods to bring organic wastes to combustion temperatures and then destroys remaiining organics in a gas-fired chamber.[4] The commercial unit is transportable and allows for discontinuous operation. Test results based on ten case studies, which include two full-scale applications at Superfund sites, show that the system can meet RCRA and TSCA incineration standards. Most of the metals in the feed concentrate in the ash. Available data suggest that this process is suitable for solid wastes containing particles

from 5 microns to 2 inches in diameter, up to 50 percent moisture content, up to 5 wt. percent chlorine and sulfur, up to 1 wt. percent alkali metals, and heating values up to 10,000 Btu/lb.[5] Based on recent commercial operations, on-line reliability should range from 50 to 75 percent. Excluding waste excavation, feed preparation, profit, and ash disposal costs, costs may be as high as $800/ton. In 1990, three vendors owned and operated commercial, full-scale infrared thermal treatment systems.[6]

American Combustion Technologies developed the Pyretron Oxygen Air Burner system, which employs an oxygen-air-fuel burner for use with a rotary kiln incinerator.[7] The replacement of air with oxygen reduces the heat sink phenomenon of nitrogen in air and reduces the combustion gas volume, which in turn reduces particulate carryover and the load on air pollution control equipment. The primary advantage of the burner system is increased waste throughput, which, based primarily on SITE data, can be twice that of a conventional rotary kiln incinerator. The greatest capacity increases will occur for wastes with a low heating value. During the SITE demonstration, equipment operation was fairly reliable. Representative cost savings over conventional rotary kiln incineration are estimated to be about $30 per ton. Actual savings will depend on site-specific conditions.

Terra-Vac, Inc., developed the In-Situ Vacuum Extraction process, which removes volatile organic constituents (VOCs) from the unsaturated zone of soils through extraction wells.[8] The extracted gases and water proceed to a vapor–liquid separator and an emission control system. The process has been employed at over 60 sites. Four case studies, including three Superfund sites, have been documented.[9] The process represents a viable technology to fully remediate sites contaminated with volatile organics. Considerations for use include contaminent volatility, site-specific cleanup level, and soil properties. The process works well with most soil types. The air-filled porosity of a soil is an important criterion to indicate whether vacuum extraction will work. Soils with low permeability but with adequate air-filled porosity are amenable to treatment by this process. Soils with a high water content may require extensive dewatering before the process being to work. The in situ vacuum extraction process is relatively simple and operates reliably under all weather conditions. It typically costs around $50 per ton, with the cost strongly dependent on whether off-gas treatment is required and whether wastewater is generated by the process at a given site.

International Waste Technologies and Geo-Con, Inc. together demonstrated the In-Situ Soil Stabilization system on PCB-contaminated soil. The system employs a proprietary additive and a specific soil-mixing technology which precludes the need for soil excavation. A hollow-stem auger, offered by Geo-Con, injected the IWT chemicals into the soil and mixed it. Physical properties of the treated soil were generally satisfactory: unconfined compressive strength was high, permeability was low, and little degradation occurred as a result of wet/dry weathering tests. However, many freeze/thaw samples showed marked degradation. Microstructural analysis showed a dense, homogeneous, and low-porosity structure. Overall, the results also indicated a potential for long-term durability. The process appears to immobilize PCBs; however, results near detection limits did not allow a decisive determination. The Geo-Con equipment can operate in virtually all types of soils, although clays, oily sands, and cohesive soil may reduce the treatment rate and the depth. Also, small changes in ground level result from the addition of the reagents. The Geo-Con equipment operated satisfactorily during demonstration runs, and the process is economical. The cost of a commercial, four-auger operations was estimated at $110 per ton.[10]

C.F. Systems Corporation's organic solvent extraction technology was pilot-tested in September 1988 on polychlorinated biphenyl (PCB)–contaminated harbor sediments from the New Bedford, MA Superfund site. The technology extracts organics from contaminated soils based on the solubility of organics in a mixture of liquefied propane and butane.

Sediments were treated with C.F. Systems' trailor-mounted Pit Cleanup Unit (PCU) that has a design capacity of 1.5 gpm or 20 barrels/day. PCB extraction efficiencies of 90 percent were achieved for harbor sediments containing from 350 to 2570 ppm, resulting in clean sediments containing as little as 8 ppm. The cost of applying the technology to a full-scale cleanup effort could vary from $148 and $447 per tone including pre- and post-treatment costs, materials handling, and costs for specialized factors for a full-scale unit.[11]

HAZARDOUS WASTES

General Considerations

A waste is generally thought of as any solid, liquid, or contained gaseous material that is no longer useful to its owner and is to be recycled, disposed of, or stored for later disposal. As a routine result of doing business, a company can generate many wastes, which can, if not handled or disposed of properly, cause injury or death or pollute and damage land, air, and water. Such wastes are considered hazardous and are currently regulated by numerous federal and state public health and environmental safety laws.

There are two ways in which wastes may be brought into the hazardous waste regulatory system: (1) by listing and (2) by identification through their characteristics.

Listed Wastes. A waste is considered hazardous if it appears on any one of the four lists of hazardous wastes contained in RCRA (Resource Conservation and Recovery Act) regulations. The regulations list over 400 hazardous wastes, including discarded commercial products and wastes derived from manufacturing processes.

Some wastes are considered to be "acutely hazardous," having been determined by EPA to be so dangerous in small amounts that they are regulated in the same way as are large amounts of other hazardous wastes.

Characteristic Wastes. Some wastes are considered hazardous even though they do not appear on the EPA lists because they have one or more of the following characteristics:

- They are easily combustible or flammable. Examples of such "ignitable" wastes are paint wastes, certain degreasers, and other solvents.
- They dissolve metals or other materials, or burn the skin. Examples of such "corrosive" wastes include rust removers, waste acid or alkaline cleaning fluids, and waste battery acid.
- They are unstable or undergo rapid or violent chemical reaction with water or other materials. Examples of these "reactive" wastes are waste bleaches, other waste oxidizers, and cyanide plating wastes.
- They exhibit EP (extraction procedure) toxicity. Wastes are EP-toxic if an extract from the waste is tested and found to contain high concentrations of heavy metals (mercury, lead, cadmium) or specific pesticides that could be released into the groundwater.

Hazardous Waste Defined

Defining what constitutes a "hazardous waste" requires consideration of legal and scientific factors. The basic definitions used in this chapter are derived from (a) the Resource Conservation and Recovery Act (RCRA), as amended by the Hazardous and Solid Waste Amendments (HSWA), and (b) the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA), as amended by the Superfund Amendments and Reauthorization Act (SARA). These definitions are given below.

Hazardous waste refers to "a solid waste, or combination of solid wastes, which because of its quantity, concentration, or physical, chemical, or infectious characteristics may ... pose a hazard to human health or the environment ..." [RCRA, Section 1004(5)].

Solid waste means "any garbage, refuse, sludge from a waste treatment plant, water supply treatment plant, or air pollution control facility and other discarded material,

including solid, liquid, semisolid, or contained gaseous material resulting from industrial, commercial, mining, and agricultural operations and from community activities, but does not include solid or dissolved material in domestic sewage ..." [RCRA, Section 1004(27)].

Hazardous substance means "any substances designated in Section 311 (b) (2) (A) of the Federal Water Pollution Control Act ... any hazardous waste having the characteristics identified in Section 3001 of the Solid Waste Disposal Act ... any toxic pollutant listed under Section 307(a) of the Federal Water Pollution Control Act, any hazardous air pollutant listed under Section 112 of the Clean Air Act, and any imminently hazardous chemical substance or mixture ... listed in Section 7 of the Toxic Substances Control Act" [CERCLA, Section 101(14)].

Under RCRA, a waste is considered hazardous if it is reactive, ignitable, corrosive, or toxic, or if the waste is listed as a hazardous waste in Parts 261.33-33 of the *Code of Federal Regulations* (40CFR). Currently, there are about 400 listed wastes.

In addition to hazardous wastes defined under RCRA, there are "hazardous substances" defined by Superfund.* Superfund's definition of a hazardous substance is broad and grows out of the statutory definitions in the Clean Water Act (CWA), the Clean Air Act (CAA), the Toxic Substances Control Act (TSCA), and RCRA. Superfund considers a hazardous substance to be any air or water pollutant defined as hazardous in the CAA or the CWA, or as a hazardous waste defined in RCRA, that exceeds reportable quantity levels. Superfund encompasses numerous wastes and chemical substances:

- One hundred twenty-six "priority pollutants" and 65 "toxic pollutants" (CWA).
- Sixteen "hazardous air pollutants" and 7 chemicals identified as "criteria pollutants" (CAA).
- Ninety-five chemicals defined as toxic on the basis of production volume, exposure, and biological effects (TSCA).
- Substances identified or listed as hazardous waste under §3001 of RCRA.
- CERCLA Contract Lab List.

There also are other hazardous wastes or toxic substances that are not specifically included. Examples are:

- Wastes from nonpoint sources (e.g., storm or irrigation runoff).
- Wastes that are exempt from RCRA management, such as wastewaters treated in enclosed tanks and discharges subject to NPDES permits.
- Industrial wastewaters discharges to surface waters under NPDES permits or into underground injection systems.
- Wastes from agricultural application of pesticides.
- PCB wastes that are not contaminated with hazardous waste.

Although these wastes are not regulated under RCRA or Superfund, they are covered by other statutory authorities such as the Clean Air Act, the Clean Water Act, the Safe Drinking Water Act, and the Toxic Substances Control Act.

Hazardous Waste Management Overview

The following discussion examines hazardous waste issues from an integrated perspective suitable for addressing various environmental issues; namely:

1. It gives an overview of the hazardous waste system that can serve as a point of reference for continuing, more detailed work.
2. It serves as a starting point for the 20-year state capacity certification efforts required by the Superfund Amendments and Reauthorization Act of 1986 (SARA).
3. It provides an initial vehicle for discussion of possible impacts of various EPA regulatory efforts now under way regarding hazardous waste control.

*"Superfund" became the common/popular name for CERCLA. Its goal was to clean up abandoned hazardous waste sites. When CERCLA expired, it was reauthorized and expanded under SARA.

The primary focus is on hazardous waste as defined in RCRA and by the Superfund program. However, the hazardous waste system does not exist in isolation. Before dealing with the specific subject of hazardous wastes, it is useful to look at several key elements of the overall waste system. Such an overview is provided in Fig. 3.11. The point of this diagram is to illustrate the relationships among such items as hazardous wastes, municipal and industrial wastewater, and nonhazardous solid wastes. It will be increasingly important to understand "where things go" in dealing with hazardous and solid wastes and wastewaters.

It is also important to note that this overview does not deal with every source or possible exposure route. For example, hazardous air emissions, pesticide applications, and many nonpoint waste sources are not

Fig. 3.11. Simplified water system chart. (*Courtesy EPA.*)

TABLE 3.11 Volumes of RCRA Hazardous Waste Treated and Disposed

	On-site at Private Firms (MMT)	*Off-site at Commercial Firms (MMT)*	*Estimated Total (MMT)*
Incineration	1.7	0.4	2.1
Other treatment	204.0	1.6	205.6
Solvent and other recovery	*	*	57.0
Furnaces and boilers	*	*	0.9
Land disposal	8.0	5.0	13.0
Deep-well injection	*	*	~25.0
Total			303.6**

Source: National Screening Survey, U.S. EPA, Office of Solid Waste (1986).
*Breakdown of on-site versus off-site waste volumes is not currently available.
**Total exceeds 275 MMT as deep-well injection, which is covered under the Safe Drinking Water Act, is included.
MMT = million metric wet tons per year.

included in Fig. 3.11. These sources and others such as waste discharges to surface waters through a National Pollutant Discharge Elimination System (NPDES) permit are not discussed in detail in this section.

RCRA and Superfund hazardous wastes from various industries are shown in Table 3.7, which includes the amount of waste, the number of facilities, and a general description. The hazardous wastes managed under RCRA totaled about 275 million metric tons in 1985.* The overwhelming portion of this quantity, well over 99 percent, was managed by "large quantity generators" (over 1000 kilograms per month).

A number of factors could affect the volumes of wastes generated in the future. Factors such as economics and population growth, regulatory decisions, and waste minimization efforts will influence the amount of waste produced. As industrial production increases, the volume of industrial wastes produced could also increase. However, efforts to recycle and reuse wastes, as well as programs to minimize the amount of waste generated, are likely to temper increases in waste volumes.

*Source: National screening survey conducted by the U.S. EPA Office of Solid Waste.

Hazardous Waste Treatment and Disposal

About 96 percent of all RCRA hazardous wastes are managed on-site by the companies that generate them. The remaining 4 percent goes to off-site commercial treatment and disposal facilities. Table 3.11 provides an overview of the kinds and quantities of hazardous wastes managed by 3000 facilities.

Several different kinds of treatment and disposal technologies are used in managing hazardous wastes. For example, there were over 175 on-site incinerators and 15 commercial incinerators in operation in the late 1980s. In addition to incineration, many other treatment methods are used, including biological wastewater treatment, solidification, steam stripping, and treatment impoundments. Most current systems consist of on-site impoundments handling relatively dilute wastewaters. There are about 430 operating land disposal facilities, approximately 60 of which are commercial facilities that accept a wide range of wastes. Although the majority of deep-well injection systems are located on-site, a few commercial facilities also use deep-well injection systems.

The Hazardous and Solid Waste Amendments of 1984 (HSWA) require that EPA ban the land disposal of over 400 chemicals and waste streams unless the waste streams are

TABLE 3.12 Applicability of Treatment and Disposal Technologies to Categories of Hazardous Waste

	Organic Chemicals		*Hazardous Waste Category*			
	Biodegradable	*Nonbiodegradable*	*Metals Inorganics*	*Nonmetal Caistocs*	*Acid Treatment*	*Pesticides*
Treatment	Dissolved air flotation Aerated lagoons Trickling filter Waste stab. Activated slud. Anaerobic dig. Hydrolysis	Evaporation Distillation Oil/water sep. Reverse osm. Ultrafilt. Hydrolysis Carbon adsorp. sed.	Precip/floc/ sed. Reverse osm. Ultrafilt. Chem. oxid./ red.	Precip/floc/ sed. Filtration Evaporation Chem. oxid./ red. Hydrolysis	Chem. oxid./ red.	Hydrolysis
Solidification	Chemical fixation Encapsulation					
Disposal	Incineration Land disposal Evaporation pond					

Source: SCS Engineers, 11260 Roger Bacon Drive, Reston, VA 22090.

treated, or it can be demonstrated that there will be "no migration as long as the waste remains hazardous." The solvents and dioxin portion of the rules became effective in November 1986, and other rules were promulgated over succeeding years. The major impact of these rules was to significantly increase the degree of treatment required for hazardous wastes.

Treatment Technology, Costs, and Potential risks

This section concerns principally the data and information contained in several tables and figures. Table 3.12 shows the applicability of various treatment and disposal technologies to different hazardous waste categories and is self-explanatory. Table 3.13 contains information on the costs associated with common landfilling of hazardous wastes. It is interesting to note the wide spectrum of costs associated with this simple technology, depending on the hazardous nature of the waste and the form in which it is disposed. Stringent regulations apply to the design and use of hazardous waste landfills, and carefully engineered liners and caps are required. Figure 3.12 shows the

TABLE 3.13 Range of Costs Associated with Landfilling Hazardous Waste

Process	*Cost**
Shipping to out-of-state landfills	\$100–\$300/ton
Landfilling dioxin-contaminated soils	\$70/ton
Landfilling materials with PCB concentrations < 500 ppm	\$0.50–\$1.25/gal
Landfilling low-hazard waste	\$25–\$50/ton
Landfilling high-hazard waste	\$100–\$300/ton
Landfilling drums	\$140–\$230/ton
Landfilling bulk	\$33–\$90/ton

*Source: CDM–FPC Inc. Report to EPA 1989 Superfund (in 1988 dollars).

TABLE 3.14 Range of Costs Associated with Liquid Hazardous Waste Treatment

Activated sludge	\$100,000–\$390,000/mgd
Lagoons	\$80,000–3.4 million/mgd
Rotating biological contractors	\$0.9–\$29.6 million/mgd
Carbon treatment	\$14,132–\$643.000/mgd
Oil/water separator	\$12,720/ton
Deep well injection	\$29–\$120/ton \$0.12–\$0.50/gal
Solvent recovery	\$300–\$624/ton \$1.25–\$2.60/gal
Air stripping	\$607,000–\$7.3 million/mg

*From Federal Programs Corp.–CDM report to EPA 1989 Superfund (based on 1988 dollars).

Fig. 3.12. RCRA cap and double liner. (*Land Disposal of Hazardous Waste: Proceedings of 11th Annual Research Symposium, July 1985; courtesy EPA.*)

TABLE 3.15 Range of Costs Associated with Hazardous Waste Cleanup. (In 1988 dollars.)

Surface sealing	\$1.32–\$16.88/sq yd
Grading	\$4,000–\$16,205/acre
Drainage ditches	\$1.27–\$6.04/linear foot
Revegetation	\$1.214–\$8,000/acre
Permeable treatment beds	\$14–\$267/sq ft
Deep well system	\$4,862–\$13,513/well
Extraction/injection well system	\$37.50/vertical foot
Extract wells/seepage basin	\$33,618–\$53,360/system
Excavation, transport and disposal	\$379–\$343/cu yd
Hydraulic dredging	\$1.25–\$3.54/cu yd
Mechanical dredging	\$1.37–\$0.49/cu yd
Drum handling	\$60–\$1,528/drum
Slurry wall construction	\$250,000–\$470,000 +/−50%

*From Federal Programs Corp.–CDM Report to EPA 1989 Superfund.

details of an RCRA-approved double liner placed over a suitable clay soil. Table 3.14 contains information on the costs associated with various treatment options for liquid hazardous wastes, and Table 3.15 shows the cost for hazardous waste cleanup. The latter table is particularly interesting in that it displays the wide range of technologies associated with cleanup operations.

It must be realized that there are potential environmental risks associated with all hazardous waste treatment/disposal alternatives. Some of these are listed in Table 3.16, which compares various technologies with potential impacts on health, surface water pollution, subsurface pollution, air emissions, and production of residual concentrates.

TABLE 3.16 Potential Environmental Risks Associated with Hazardous Waste Treatment/Disposal Alternatives

Technology	*Potential for Health Impacts*	*Potential for Surface Water Pollution*	*Potential for Subsurface Pollution*	*Potential for Air Emissions*	*Ash/Sludge/ Concentrate Production*
Precipitation/flocculation/ sedimentation	0	+	−	−	+
Filtration	0	+	−	−	+
Evaporation	0	0	−	+	+
Distillation	0	0	−	+	+
Flotation	0	+	−	0	+
Reverse osmosis	0	−	−	−	+
Ultrafiltration	0	−	−	−	+
Chemical oxidation/reduction	0	+	−	−	0
Hydrolysis	0	0	−	−	0
Aerated lagoon	+	+	+	+	−
Trickling filter	+	+	0	+	+
Waste stabilization pond	+	+	+	+	−
Anaerobic digestion	0	0	0	0	+
Activated sludge	0	+	0	0	+
Carbon adsorption	0	+	−	−	0
Incineration	+	−	−	+	+
Land disposal	+	+	+	+	−
Chemical fixation	−	−	0	−	−
Encapsulation	−	−	0	−	−
Evaporation pond	+	+	+	+	+

+ = Possible Impact. 0 = Variable. − = No possible Impact.
Source: SCS Engineers, 11260 Roger Bacon Drive, Reston, VA 22090.

RECYCLE–REUSE

Before proceeding with this discussion, it is necessary to define the terms recycle, multiple use, and reuse. Recycle as used here means using water over and over for the same application. An examplenwould be in the use of steam to drive a turbine where the condensate is returned to the boiler. Reuse or multiple use means using water more than once, but always for a different and successively dirtier application until it is no longer fit for use. The objective of all these methods of water use is to reduce water intake and the discharge of wastewater.

In the plant depicted in Fig. 3.13, the three major water uses are in processing, cooling, and steam generation. Assumed in constructing this hypothetical situation is the suitability of the process waste water for cooling functions, as well as the acceptability of cooling water blowdown for making boiler feed. The average gross water use is taken as 15 units. In the once-through system, the water intake is obviously the same as the gross water use, that is, 15 units. In the case of multiple use and reuse, although the gross water use remains the same, the intake and discharge requirements are less—how much less depending on the technology used.

In order to reduce wastewater discharge through reuse or recycle, it is necessary continually to remove waste materials that accumulate in the water. In practice, it is the accumulation of wastes or impurities that limits the extent to which reuse is feasible. Such factors as salinity, alkalinity, organic matter, and suspended solids must be controlled so that the process in which the water is used is not compromised. In the example being discussed, the boiler feed water is made acceptable by evaporation, the alkalinity is controlled by addition of chemicals, and organic and suspended solids are removed by traditional methods. If the ratio of gross water

Fig. 3.13. Reuse-recycle industrial water system. (Lacy, W. J., "The Closed-Loop Cycle for Industrial Wastewater: The Future Pollution Solution," *Environ-ment International*, 2, 3–8 (1979)).

to water intake is an index of reuse, the once-through system would have a reuse ratio of 1.0, the multiple use system in the example would have a ratio of 1.5, and the reuse–recycle system would have a ratio of 15.0. If total reuse could be achieved—a practical impossibility because of losses from evaporation, leaks, etc.—the ratio would be infinite. It will always be necessary to provide some makeup water to the system.

INDUSTRIAL WATER USE

The total water intake by U.S. industries in 1983, according to a U.S. Department of Commerce survey, was more than 10 trillion gallons (U.S. Department of Commerce, 1986). This represented 96 percent of the total water use estimated for manufacturers. Those companies reporting water intake of 20 million gallons or more represent only 3 percent of the total manufacturing establishments, 32 percent of the total number of persons employed in manufacturing, and 46 percent of the value added by manufacturers.

Approximately 84 percent of the 10 trillion gallons of water used by manufacturing establishments in 1983 was accounted for by four major industry groups: SIC 28, Chemicals and Allied Products; SIC 33, Primary Metals; SIC 26, Paper and Allied Products; and SIC 29, Petroleum and Coal Products. Of the total water intake for manufacturing, 60 percent was used for cooling and condensing operations, such as steam for electric power generation, air conditioning, and other uses. Approximately 59 percent of the total water intake was from company surface water systems. Of the total 8.9 trillion gallons of water discharged, 45 percent was treated.

Approximately 75 percent of the water used in manufacturing industries was concentrated in six water resource regions: Great Lakes, South Atlantic–Gulf, Middle Atlantic, Ohio, Lower Mississippi, and Texas–Gulf.

Industrial Water Quality

Water use and needs in the manufacturing industries can be divided among three major functions: cooling; process use; and use in boiler feedwater for steam production. Table 3.17 shows water intake by industry, source, and kind, for all manufacturing and for selected industrial groups. These are the requirements for intake and consumption and

TABLE 3.17 Water Intake by Source and Kind for Industry Groups for 1983 (in billion gallons)

		Source						Kind		
			Public	Company						
SIC Code	Industry	Total	Water System	Surface	Ground	Tidewater	Other	Fresh	Brackish*	Salt**
	All manufacturing	10,039	1,311	5,927	1,234	1,199	369	8,522	739	777
20	Food & Kindred Products	648	219	129	232	37	31	568	62	18
26	Paper & Allied Products	1,900	257	1,212	350	62	18	1,766	115	19
28	Chemical & Allied Prod.	3,400	211	2,162	314	592	123	2,665	202	539
29	Petroleum & Coal Prod.	818	138	286	62	297	35	491	223	104
33	Primary Metals	2,363	109	1,821	149	[D]	[D]	2,211	[D]	[D]

(1) U.S. Department of Commerce, 1982 Census of Manufacturers Water Use, MC 82-5-6 (Mar. 1986).
(2) Standard Industrial Classification.
*1000 to 10,000 ppm dissolved salts.
**More than 100 ppm dissolved salts.
[D] Withheld to avoid disclosing data for individual companies; data are included in higher level totals.

TABLE 3.18 Industrial Water Quality in Use[a]

			Process and General Industrial Water						
			Iron and Steel			Pulp and Paper			Boiler Feedwater (150 psig)
Water Quality Charracteristics (ppm)	Cooling Water (once thru) Fresh	Cooling Water (once thru) Brackish	Quench and Gas Clean	DM Rinse	Chemical and Allied Products	Kraft Unbleached	Fine	Petro Refin.	Prior to Chem Cond.
Temperature, °F	(b)	(b)	100	100	(b)	—	—	—	(b)
pH (units)	5–8.3	6–8.3	5–9	—	5–5.9	—	—	6–9	7–10
S.S.	5,000	2,500	100	0.1	10,000	100 (c)	10 (c)	5,000	10
D.S. (e)	1,000	35,000	(b)	0.1	2,500	500	200	3,600	700
Chloride	600	19,000	(b)	0.1	500	—	—	1,600	(b)
Sulfate	630	2,700	—	—	850	—	—	900	(b)
Silica	50	25	—	—	(b)	100	20	85	30
Iron	(b)	(b)	—	—	10	1	0.1	15	1
Hardness	850	6,260	(b)	0.1	1,000	200	100	900	360
Alkalinity	500	115	(b)	0.5	500	150	75	500	360
C.O.D.	75	75	—	—	(b)	—	—	1000	5
Oil	(f)	(f)	(b)	0.02	—	—	—	—	—
Color (units)	—	—	—	—	500	100	5	25	—
D.O.	some	some	some	some	(b)	—	—	—	2.5

Notes: (a) Reference: Based on EDA former water quality criteria reports; (b) accepted as received; (c) suspended solids; (d) as turbidity; (e) dissolved solids; (f) no floating oil.

reflect the extent of internal water reuse that occurs within the industry. Cooling water needs are largest, followed by process demands and boiler feed makeup. This is a fortunate situation, as the water-quality requirements are generally less critical for cooling than for process or boiler feed.

Table 3.18 summarizes water-quality characteristics of the various water functions for major industrial users. It shows that industry has a use for low-quality water and could use waters of various qualitites. It further shows that if one excludes the bacterial consideration, then the quality criteria parameters are less

TABLE 3.19 Summary of Water Use Statistics for Industry Groups and Industries: 1983

SIC Code	Industry group and industry	Establishments reporting water intake of 20 million gallons or more during 1982	Gross water used[1]: Total: Quantity (billion gallons)	Gross water used[1]: Total: Percent of all manufacturing industries	Gross water used[1]: Water intake: Quantity (billion gallons)	Gross water used[1]: Water intake: Percent of all manufacturing industries	Gross water used[1]: Water and reused (billion gallons)	Water discharged: Total (billion gallons)	Water discharged: Untreated (billion gallons)	Water discharged: Treated (billion gallons)
	All manufacturing industries[2]	10,262	33,835.2	100	10,038.9	100	23,796.3	8,913.7	4,889.8	4,023.9
20	Food and kindred products	2,656	1,406.2	4	647.7	6	758.6	552.0	355.8	196.1
26	Paper and allied products	600	7,435.8	22	1,899.3	19	5,536.5	1,768.1	479.0	1,289.1
28	Chemicals and allied products	1,315	9,630.1	28	3,400.7	34	6,229.4	2,979.8	1,996.3	983.6
29	Petroleum and coal products	260	6,177.3	18	818.4	8	5,358.9	699.3	323.4	375.9
33	Primary metal industries	776	5,885.2	17	2,362.5	24	3,522.8	2,112.0	1,227.9	884.1

*U.S. Department of Commerce, 1982 Census of Manufacturers Water Use, MC 82-5-6 (Mar. 1986).

stringent than those for desirable potable water. These are important considerations in planning an industrial water management system, as they would allow the reuse of water in a countercurrent manner for successively dirtier applications. Combining reuse with a water recycle system can close an industrial water circuit to the point where there is essentially zero discharge of pollutants. Table 3.19 is a summary of the water discharged, treated and untreated, by point of discharge for various industrial groups.

Rather than discuss economics with typical costs based on the amount of water treated or on the specific processes employed, it may be simpler to look at the total water treated, amount of suspended solids, alkalinity, hardness, organic matter, and dissolved salts removed.

In poor-water-quality areas the total water recycle method may provide a minimum cost for pollution control, as these water supplies bring into a plant several of the same materials whose removal is required for use or discharge. The net effect is that more suspended solids, organic materials, and dissolved salts would be handled on a once-through basis than if only net increases due to internal (closed-loop) plant functions were to be treated.

In 1990, manufacturing plants spent over $6 billion for pollution abatement capital expenditures, according to the Bureau of the Census, U.S. Department of Commerce. Of this total, about $2.5 billion was spent for air pollution abatement, $2.6 billion for water pollution abatement, and $817 million for solid waste management. In addition, industrial firms spent about $17 billion for pollution abatement operating costs. A breakdown by different industry groups is given in Table 3.20. A graph showing expenditures for pollution abatement and control over the period from 1972 to 1989 is shown in Fig. 3.14.

Research and Development Needs

Much research and development is needed to achieve a realistic zero discharge of water pollutants for the wide variety of industrial

TABLE 3.20 Pollution Abatement Capital Expenditures and Operating Costs, by Form of Abatement and major Industry Group: 1990

(Value figures in millions of dollars)

SIC Code	Industry	Pollution abatement capital expenditures (PAGE)				Pollution abatement gross annual cost (GAC) including payment to government units				Standard error of estimate (percent)	
		Total	Air	Water	Solid waste	Total	Air	Water	Solid waste	PAGE	GAC
	All industries	6,029	2,562	2,650	817	17,104	5,027	6,414	5,663	2	2
20	Food	249	65	163	21	1,109	146	693	271	11	4
21	Tobacco	6	4	1	1	45	16	13	16	30	17
22	Textile	46	19	25	2	198	26	115	58	25	8
24	Lumber	105	58	10	37	259	86	37	136	25	12
25	Furniture	24	19	2	2	141	51	23	67	35	16
26	Paper	1,075	414	510	152	1,607	398	788	421	3	4
27	Printing	68	56	4	7	241	71	31	140	23	10
28	Chemicals	1,851	596	994	261	3,948	858	1,797	1,293	4	2
29	Petroleum	917	426	401	90	2,705	1,472	702	531	3	4
30	Rubber	94	69	11	14	428	97	113	218	15	10
31	Leather	8	1	7	1	49	7	26	16	25	25
32	Stone	128	94	20	14	498	248	93	157	10	5
33	Primary metals	499	279	167	54	2,054	944	565	545	4	3
34	Fabricated metals	171	53	61	57	814	114	344	356	9	8
35	Machinery	107	46	41	20	577	78	159	320	10	7
36	Electric	178	93	58	26	788	113	370	304	12	14
37	Transportation	395	207	143	46	1,232	247	373	612	10	4
38	Instruments	92	58	26	8	248	37	69	142	10	7
39	Miscellaneous	18	7	7	4	185	19	103	63	19	59

Note: Totals may not agree with detail because of independent rounding.
Source: Government Printing Office: 1992 – 311-899/40231

water problems. If the approach taken to zero discharge of pollutants is related to integrated supply and wastewater management, major areas for R and D can be identified, as shown in Fig. 3.15. The figure also shows various ways to reuse–recycle water and to manage sludges, and general areas where research and development may lead to a better way to reduce water pollution outside an industrial plant.

The major industrial water use and handling activities (i.e., cooling, process operations, water treatment, and steam production) should be viewed individually for opportunities to improve pollution control performance and reduce pollutants, as well as in an integrated fashion to conserve energy and avoid duplication in treatment. For example, in process operations improved housekeeping, revised operating procedures, and process changes in themselves can produce significant gains in reducing water management costs while improving pollution control. As most industrial water is used for cooling purposes, research and development efforts in cooling techniques, water quality limitations, corrosion control, and the effectiveness of the cooling process itself as a method of treating pollutants (other than heat) should be strongly pursued.

The challenge of achieving zero discharge of pollutants from the major water use generally is best left to the chemical engineering profession. The problems of the four large industrial water users are not sanitary in the classical sense but instead are chemical, or, more specifically, they involve chemical separation of pollutants in water to a desired level that will allow reuse or discharge.

Fig. 3.14. Real expenditures for pollution abatement and control, 1972–89. (Rutledge, G. L. and Leonard, M. L., U.S. Department of Commerce, Bureau of Economic Analysis.) (Note: Pollution Abatement and Control (PAC) expenditures consist of expenditures for the direct abatement of pollution (PA), for regulation and monitoring, and for research and development.)

Another problem is that of optimizing and integrating water conditioning and treatment steps to achieve a minimum number of operations.

The system described here also includes consideration of environmental concerns other than water pollution control, namely, the use of wastewater for air and thermal pollution control functions and waste residue usage for heat and power production and/or water treatment (ash and sludges). However, even an ideal plant system will have a net discharge of some waste material. The system described in Fig. 3.14 includes excess ash from thermal power production and wastewater blowdown of high salinity, hardness, and toxicity. These wastes may be blended for some beneficial purpose, to ease proper handling, or for controlled assimilation by the environment in appropriate disposal sites. These concerns will have to be given consideration as industry proceeds to meet future environmental demands.

The industrial water management concept of closed-cycle operation to achieve zero discharge of pollutants appears to be a

Fig. 3.15. Industrial wastewater management R and D needs.

historically, economically, and theoretically sound approach to industrial wastewater pollution control. Comments about excessive or unreasonable energy demands for industrial water pollution control may be more conjecture than fact. It is difficult to conceive how a water conservation approach to pollution control, properly developed, would not also be energy-conserving.

Future Demands

Closed-loop industrial wastewater and water systems are vitally necessary to maintain continuity in future industrial expansion. The huge water demands and high water usage growth rate of U.S. industry cannot continue to rely on traditional water supply sources. Even in water-abundant areas, intake water supplies for industrial use are fast becoming restricted.

Current and future environmental standards for wastewater discharges are expected to increase the pressure on industry to reduce both the pollution discharge loads and the magnitude of effluent volumes in order to minimize environmental impact. Industrial water-quality requirements for reuse are less demanding, as a general rule, than are municipal requirements. Accordingly, industrial water reuse should be technically and economically achievable earlier than comparable municipal water reuse systems.

Wastewater reuse is not only a resource conservation measure, but also a method of pollution control—a step in tune with future demands. Adequate research and development activity in this area is crucial to accelerating the implementation of extensive wastewater reuse systems and, eventually, the total closed-loop cycle. The latter, with no effluent discharge, would comply with any quality standards, now or in the future.

TERMINOLOGY

Many of the more common terms encountered in wastewater-treatment technology are given below to assist the reader in developing a vocabulary in this field and in understanding what follows.

Abatement. The method of reducing the degree of intensity of pollution; also the use of such a method.

Absorption. The penetration of a substance into or through another; for example, in air pollution control, absorption in a liquid from which it then can be extracted.

Acclimation. The physiological and behaviorial adjustments of an organism to changes in its immediate environment.

Acclimatization. The acclimation or adaptation of a particular species over several generations to a marked change in the environment.

Activated carbon. A highly adsorbent form of carbon, used to remove odors and toxic substances from gaseous emissions. In advanced waste treatment, activated carbon is used to remove dissolved organic matter from wastewater.

Activated sludge. Sludge that has been aerated and subject to bacterial action, used to remove organic matter from sewage.

Activated-sludge process. The process of using biologically active sewage sludge to hasten breakdown of organic matter in raw sewage during secondary waste treatment.

Acute toxicity. Any poisonous effect produced within a short period of time, usually from 24 to 96 hours, resulting in severe biological harm and often death.

Adaption. A change in structure or habit of an organism that allows it to exist in a changed environment.

Adsorption. The adhesion of a substance to the surface of a solid or a liquid. Adsorption often is used to extract pollutants by causing them to be attached to such adsorbents as activated carbon or silica gel. Hydrophobic, or water-repulsing, adsorbents are used to extract oil from waterways in oil spills.

Advanced waste treatment. Wastewater treatment beyond the secondary or biological stage that includes removal of

nutrients such as phosphorus and nitrogen and removal of a high percentage of suspended solids. Advanced waste treatment, known as tertiary treatment, is the "polishing stage" of wastewater treatment and produces a high-quality effluent.

Aeration. The process of being supplied or impregnated with air. Aeration is used in wastewater treatment to foster biological and chemical purification.

Aerobic. Term referring to life or processes that can occur only in the presence of oxygen.

Agricultural pollution. The liquid and solid wastes from all types of framing, including runoff from pesticides, fertilizers, and feedlots; water and wind erosion and dust from plowing; animal manure; carcasses; and crop residues and debris. It has been estimated that agricultural pollution in the United States has amounted to more than $2\frac{1}{2}$ billion tons per year.

Air stripping. A treatment system that removes, or "strips," volatile organic compounds (VOCs) from contaminated groundwater of surface water by forcing an airstream through the water and causing the compounds to evaporate.

Anaerobic. Term referring to life or processes that occur in the absence of oxygen.

Anti-degradation clause. A provision in air quality and water quality laws that prohibits deterioration of air or water quality in areas where the pollution levels are presently below those allowed.

Aquaculture project. A controlled discharge of pollutants to enhance growth or propagation of harvestable freshwater, estuarine, or marine plant or animal species.

Aquatic plants. Plants that grow in water, either floating on the surface, growing up from the bottom of the body of water, or growing under the surface of the water.

Aquifer. An underground bed or stratum of gravel or porous stone that contains water.

Assimilation. Conversion or incorporation of absorbed nutrients into protoplasm. The term also refers to the ability of a body of water to purify itself of organic pollution.

Autotrophic. Self-nourishing; term denoting those organisms capable of constructing organic matter from inorganic substances.

Bacteria. Single-celled microorganisms that lack chlorophyll. Some bacteria are capable of causing human, animal, or plant diseases; others are essential in pollution control because they break down organic matter on the air and in the water.

Baghouse. An air pollution abatement device used to trap particulates by filtering gas streams through large fabric bags, usually made of glass fibers.

Baling. A means of reducing the volume of solid waste by compaction.

Bar screen. In wastewater treatment, a screen that removes large floating and suspended solids.

Benthic region. The bottom of a body of water. This region supports the benthos, a type of life that not only lives upon but contributes to the character of the bottom.

Benthos. The plant and animal life whose habitat is the bottom of a sea, lake, or river.

Bioassay. The employment of living organisms to determine the biological effect of some substance, factor, or condition.

Biochemical oxygen demand (BOD). A measure of the amount of oxygen consumed in the biological processes that break down organic matter in water. Large amounts of organic waste use up large amounts of dissolved oxygen; thus the greater the degree of pollution is, the greater the BOD.

Biodegradable. Capable of decomposing quickly as a result of the action of microorganisms.

Biological oxidation. The process by which bacterial and other microorganisms feed on complex organic materials and decompose

them. Self-purification of waterways, activated sludge, and trickling filter wastewater treatment processes depend on this principle. The process is also called biochemical oxidation.

BOD_5. The amount of dissolved oxygen consumed in 5 days by biological processes breaking down organic matter in an effluent. (See *Biochemical oxygen demand.*)

Boom. A floating device that is used to contain oil on a body of water.

Carcinogen. A substance that causes cancer.

Carbon adsorption. A treatment system where contaminants are removed from groundwater or surface water when the water is forced through tanks containing activated carbon, a specially treated material that attracts the contaminants.

Channelization. The straightening and deepening of streams to permit water to move faster, to reduce flooding, or to drain marshy acreage for farming. Channelization also reduces the organic waste assimilation capacity of the stream and may disturb fish breeding and destroy the stream's natural beauty.

Chemical oxygen demand (COD). A measure of the amount of oxygen required to oxidize organic and oxidizable inorganic compounds in water. The COD test, like the BOD test, is used to determine the degree of pollution in an effluent.

Chlorination. The application of chlorine to drinking water, sewage, or industrial waste for disinfection or oxidation of undesirable compounds.

Chlorinator. A device for adding a chlorine-containing gas or liquid to drinking water or wastewater.

Chlorine-contact chamber. A chamber in a waste treatment plant in which effluent is disinfected by chlorine before it is discharged to the receiving waters.

Clarification. In wastewater treatment, the removal of turbidity and suspended solids by settling, often aided by centrifugal action and chemically induced coagulation.

Clarifier. In wastewater treatment, a settling tank that mechanically removes settleable solids from wastes.

Coagulation. The clumping of particles in order to settle out impurities; often induced by chemicals such as lime or alum.

Coliform index. An index of the purity of water based on a count of its coliform bacteria.

Coliform organism. Any of a number of organisms common to the intestinal tract of humans and animals whose presence in wastewater is an indicator of pollution and of potentially dangerous bacterial contamination.

Combined sewers. A sewerage system that carries both sanitary sewage and stormwater runoff. During dry weather, combined sewers carry all wastewater to the treatment plant. During a storm, only part of the flow is intercepted, because of plant overloading; the remainder goes untreated to the receiving stream.

Combustion. Burning; technically, a rapid oxidation accompanied by the release of energy in the form of heat and light. It is one of three basic contributing factors causing air pollution; the others are attrition and vaporization.

Comminution. Mechanical shredding or pulverizing of waste; a process that converts waste into homogeneous and more manageable material. It is used in solid waste management and in the primary stage of wastewater treatment.

Comminutor. A device that grinds solids to make them easier to treat.

Compaction. Reducing the bulk of solid waste by rolling and tamping.

Compost. Relatively stable decomposed organic material.

Composting. A controlled process of degrading organic matter by microorganisms.

(1) Mechanical—a method in which the compost is continuously and mechanically mixed and aerated. (2) Ventilated cell—method where compost is mixed and aerated by being dropped through a vertical series of ventilated cells. (3) Windrow—an open-air method in which compostable material is placed in windrows, piles, or ventilated bins or pits and occasionally turned or mixed. The process may be anaerobic or aerobic.

Decomposition. Reduction of the net energy level and a change in chemical composition of organic matter brought about by the actions of aerobic or anaerobic microorganisms.

Diatomaceous earth (Diatomite). A fine siliceous material resembling chalk used in wastewater treatment plants to filter sewage effluent in order to remove solids. It also may be used as an inactive ingredient in pesticide formulations applied as dust or powder.

Digester. In a wastewater treatment plant, a closed tank that decreases the volume of solids and stabilized raw sludge by bacterial action.

Digestion. The biochemical decomposition of organic matter. Digestion of sewage sludge takes place in tanks where the sludge decomposes, resulting in partial gasification, liquefaction, and mineralization of pollutants.

Dissolved oxygen (DO). The oxygen dissolved in water or sewage. Adequately dissolved oxygen is necessary for the life of fish and other aquatic organisms and for the prevention of offensive odors. Low dissolved oxygen concentrations generally are due to the discharge of excessive organic solids having a high BOD, the result of inadequate waste treatment.

Dissolved solids. The total amount of dissolved material, organic and inorganic, contained in water or wastes. Excessive dissolved solids make water unpalatable for drinking and unsuitable for industrial uses.

Effluent. A discharge of pollutants into the environment, partially or completely treated or in its natural stage. Generally the term is used in regard to discharges into waters.

Electrodialysis. A process that uses electrical current and an arrangement of permeable membranes to separate soluble minerals from water. Often it is used to desalinize salt or brackish water.

Enrichment. The addition of nitrogen, phosphorus, and carbon compounds or other nutrients into a lake or other waterway in order to greatly increase the growth potential of algae and other aquatic plants. Most frequently, enrichment results from the inflow of sewage effluent or from agricultural runoff.

Environment. The sum of all external conditions and influences affecting the life, the development, and, ultimately, the survival of an organism.

Environmental impact statement. A document prepared by a federal agency on the environmental impact of its proposals for legislation and other major actions significantly affecting the quality of the human environment. Environmental impact statements are used as tools for decision making and are required by the National Environmental Policy Act.

Eutrophication. The normally slow aging process by which a lake evolves into a bog or marsh and ultimately assumes a completely terrestrial state and disappears. During eutrophication the lake becomes so rich in nutritive compounds, especially nitrogen and phosphorus, that algae and other microscopic plant life become superabundant, thereby "choking" the lake and eventually causing it to dry up. Eutrophication may be accelerated by many human activities.

Eutrophic lakes. Shallow lakes, weed-choked at the edges and very rich in nutrients. The water is characterized by large amounts of algae, low transparency, low dissolved oxygen, and high BOD.

Evaporation ponds. Shallow, artificial ponds where sewage sludge is pumped, permitted to dry, and either removed or buried by more sludge.

Fecal coliform bacteria. A group of organisms common to the intestinal tracts of humans and of animals. The presence of fecal coliform bacteria in water is an indicator of pollution and of potentially dangerous bacterial contamination.

Filtration. In wastewater treatment, the mechanical process that removes particulate matter by separating water from solid material, usually by passing it through sand.

Floc. A clump of solids formed in sewage by biological or chemical action.

Flocculation. In wastewater treatment, the process of separating suspended solids by chemical creation of clumps or flocs.

Fly ash. All solids, including ash, charred paper, cinders, dust, soot, or other partially incinerated matter, that are carried in a gas stream.

Groundwater. The supply of freshwater under the earth's surface in an aquifer or soil that forms a natural reservoir for human use.

Groundwater runoff. Groundwater that is discharged into a stream channel as spring or seepage water.

Habitat. The sum total of environmental conditions of a specific place that is occupied by an organism, a population, or a community.

Hazardous substance. Any material that poses a threat to public health and/or the environment. Typical hazardous substances are materials that are toxic, corrosive, ignitable, explosive, or chemically reactive.

Heavy metals. Metallic elements with high molecular weights, generally toxic in low concentration to plant and animal life. Such metals are often residual in the environment and exhibit biological accumulation. Examples include mercury, chromium, cadmium, arsenic, and lead.

Hydrology. Science dealing with the properties, movement, and effects of water on the earth's surface, in the soil and rocks below, and in the atmosphere.

Incineration. Burning of certain types of solid, liquid, or gaseous materials under controlled conditions to destroy hazardous waste.

Infiltration/inflow. Total quantity of water entering a sewer system. Infiltration means entry through such sources as defective pipes, pipe joints, connections, or manhole walls. Inflow signifies discharge into the sewer system through service connections from such sources as the area of foundation drainage, springs and swamps, storm waters, street wash waters, or sewers.

Interceptor sewers. Sewers used to collect the flows from main and trunk sewers and carry them to a central point for treatment and discharge. In a combined sewer system, where street runoff from rains is allowed to enter the system along with sewage, interceptor sewers allow some of the sewage to flow untreated directly into the receiving stream, to prevent the plant from being overloaded.

Lagoon. In wastewater treatment, a shallow pond—usually artificial—where sunlight, bacterial action, and oxygen interact to restore wastewater to a reasonable state of purity.

Leachate. Liquid that has percolated through solid waste or some other medium and has extracted dissolved or suspended materials from it.

Leaching. The process by which soluble materials in the soil, such as nutrients, pesticides, chemicals, or contaminants, are washed into a lower layer of soil or are dissolved and carried away by water.

Monitoring wells. Special wells drilled at specific locations on or off a hazardous wastes site where groundwater can be

sampled at selected depths and studied to determine such things as the direction in which the groundwater flows and the types and amounts of contaminants present.

NTA. Nitrilotriacetic acid, a compound once used to replace phosphates in detergents.

Nutrients. Elements or compounds essential as raw materials for organism growth and development; for example, carbon, oxygen, nitrogen, and phosphorus.

Oligotrophic lakes. Deep lakes that have a low supply of nutrients and thus contain little organic matter. Such lakes are characterized by high water transparency and high dissolved oxygen.

Outfall. The mouth of a sewer, drain, or conduit where an effluent is discharged into the receiving waters.

Oxidant. Any oxygen-containing substance that reacts chemically in the air to produce new substances. Oxidants are the primary contributors to photochemical smog.

Oxidation. A chemical reaction in which oxygen unites or combines with other elements. Organic matter is oxidized by the action of aerobic bacteria; thus oxidation is used in wastewater treatment to break down organic wastes.

Oxidation pond. An artificial lake or pond in which organic wastes are destroyed by bacterial action. Often oxygen is bubbled through the pond to speed the process.

Packed tower. An air pollution control device in which polluted air is forced upward through a tower packed with crushed rock or wood chips while a liquid is sprayed downward on the packing material. The pollutants in the air stream either dissolve or react chemically with the liquid.

Pathogenic. Causing or capable of causing disease.

PCBs. Polychlorinated biphenyls, a group of organic compounds used in the manufacture of plastics. In the environment, PCBs exhibit many of the same characteristics as DDT and may, therefore, be confused with that pesticide. PCBs are highly toxic to aquatic life; they persist in the environment for long periods of time, and they are biologically accumulative.

Percolation. Downward flow or infiltration of water through the pores or spaces of a rock or soil.

Polyelectrolytes. Synthetic chemicals used to speed flocculation of solids in sewage.

Pretreatment. In wastewater treatment, any process used to reduce the pollution load before the wastewater is introduced into a main sewer system or before it is delivered to a treatment plant for substantial reduction of the pollution load.

Primary treatment. The first stage in wastewater treatment, in which substantially all floating of settleable solids are mechanically removed by screening and sedimentation.

Quality assurance/quality control (QA/QC). A system of procedures, checks, audits, and corrective actions used to ensure that field work and laboratory analysis during the investigation and cleanup of Superfund sites meet established standards.

Resource recovery. The process of obtaining or energy, particularly from solid wastes.

Removal action. An immediate action taken over the short term to address a release or threatened release of hazardous substances.

Resource Conservation and Recovery Act (RCRA). A federal law that established a regulatory system to track hazardous substances from the time of generation to disposal. The law requires safe and secure procedures to be used in treating, transporting, storing, and disposing of hazardous substances. RCRA is designed to prevent the creation of new, uncontrolled hazardous waste sites.

Response action. A CERCLA-authorized action at a Superfund site involving either a short-term removal action or a long-term

remedial response that may include, but is not limited to, the following activities:

- Removing hazardous materials from a site to an EPA-approved licensed hazardous waste facility for treatment, containment, or destruction.
- Containing the waste safely on-site to eliminate further problems.
- Destroying or treating the waste on-site by using incineration or other technologies.
- Identifying and removing the source of groundwater contamination and halting further movement of the contaminants.

Runoff. The portion of rainfall, melted snow, or irrigation water that flows across the ground surface and eventually is returned to streams. Runoff can pick up pollutants from the air or the land and carry them to the receiving waters.

Sanitation. Control of all the factors in humans' physical environment that exercise or can exercise a deleterious effect on their physical development, health, and survival.

Sanitary landfilling. An engineered method of solid-waste disposal on land in a manner that protects the environment. Waste is spread in thin layers, compacted to the smallest practical volume, and covered with soil at the end of each working day.

Screening. The removal of relatively coarse floating and suspended solids by straining with racks or screens.

Scrubber. An air pollution control device that uses a liquid spray to remove pollutants from a gas stream by absorption or chemical reaction. Scrubbers also reduce the temperature of the emission.

Secondary treatment. Wastewater treatment, beyond the primary stage, in which bacteria consume the organic parts of the wastes. This biochemical action is accomplished by the use of trickling filters or the activated-sludge process. Effective secondary treatment removes virtually all floating and settleable solids. Customarily, disinfection by chlorination is the final stage of the secondary treatment process.

Sedimentation tanks. In wastewater treatment, tanks where the solids are allowed to settle or to float as scum. Scum is skimmed off, and settled solids are pumped to incinerators, digesters, filters, or other means of disposal.

Septic tank. An underground tank used for the deposition of domestic wastes. Bacteria in the wastes decompose the organic matter, and the sludge settles to the bottom. The effluent then flows through drains into the ground. Sludge is pumped out at regular intervals.

Settleable solids. Bits of debris and fine matter heavy enough to settle out of wastewater.

Settling tank. In wastewater treatment, a tank or basin in which settleable solids are removed by gravity.

Sewage. The total of organic waste and wastewater generated by residential and commercial establishments.

Sewerage. The entire system of sewage collection, treatment, and disposal. The term also applies to all the effluent carried by sewers, whether it be sanitary sewage, industrial wastes, or storm-water runoff.

Sludge. The construction of solids removed from sewage during wastewater treatment. Sludge disposal is handled by incineration, dumping, or burial.

Solid waste. Useless, unwanted, or discarded material with insufficient liquid content to be free-flowing. (Also see *Waste.*) (1) Agricultural—solid waste that results from the raising and slaughtering of animals, as well as the processing of animal products and orchard and field crops. (2) Commercial—waste generated by stores, offices, and other activities that do not actually turn out a product. (3) Industrial—waste that results

from industrial processes and manufacturing. (4) Institutional—waste originating from educational, health care, and research facilities. (5) Municipal—residential and commercial solid waste generated within a community. (6) Pesticidal—the residue from the manufacturing, handling, or use of chemicals intended for killing plant and animal pests. (7) Residential—waste that normally originated in a residential environment; sometimes called domestic solid waste.

Sorption. A term including both adsorption and absorption. Sorption is basic to many processes used to remove gasous and particulate pollutants from an emission and to clean up oil spills.

Stabilization. The process of converting active organic matter in sewage sludge or solid wastes into inert, harmless material.

Suspended solids (SS). Small particles of solid pollutants in sewage that contribute to turbidity and that resist separation by conventional means. The examination of suspended solids and the BOD test constitute the main determinations for water quality performed at wastewater treatment facilities.

Synergism. The cooperative action of separate substances so that the total effect is greater than the sum of the effects of the substances acting independently.

Tertiary treatment. Wastewater treatment beyond the secondary or biological stage that includes removal of nutrients such as phosphorus and nitrogen and removal of a high percentage of suspended solids. Also known as advanced waste treatment, tertiary treatment produces a high-quality effluent.

Thermal pollution. Degradation of water quality by the introduction of a heated effluent. It is primarily a result of the discharge of cooling waters from industrial processes, particularly from electrical power generation. Even small deviations from normal water temperatures can affect aquatic life. Thermal pollution usually can be controlled by cooling towers.

Toxic pollutants. A combination of pollutants including disease-carrying agents that, after discharge and upon exposure, ingestion, inhalation, or assimilation into any organism can cause death or disease, mutations, deformities, or malfunctions in such organisms or their offspring.

Trace metals. Metals found in small quantities or traces, usually because of their insolubility.

Trickling filters. A device for the biological or secondary treatment of wastewater, consisting of a bed of rocks or stones that support bacterial growth. Sewage is trickled over the bed, enabling the bacteria to break down organic wastes.

Urban runoff. Storm water from city streets and gutters, which usually contains a great deal of litter and organic and bacterial wastes.

Waste. (Also see *Solid waste.*) (1) Bulky waste—items whose large size precludes or complicates their handling by normal collection, processing, or disposal methods. (2) Construction and demolition waste—building materials and rubble resulting from construction, remodeling, repair, or demolition operations. (3) Hazardous waste—waste that requires special handling to avoid illness or injury to persons or damage to property. (4) Special wastes—those wastes that require extraordinary management. (5) Wood pulp waste—wood or paper fiber residue resulting from a manufacturing process. (6) Yard waste—plant clippings, prunings, and other discarded material from yards and gardens (also known as yard rubbish).

Water pollution. The addition of sewage, industrial wastes, or other harmful or objectionable material to receiving water at concentrations or in sufficient quantities to result in measurable degradation of water quality.

Water quality criteria. The levels of pollutants that affect the suitability of water

for a given use. Generally, water use classification includes: public water supply, recreation, propagation of fish and other aquatic life, agriculture, and industrial use.

Water quality standard. A plan for water quality management containing four major elements: (1) the uses (recreation, drinking water, fish and wildlife propagation, industrial, or agricultural) to be made of the water, (2) criteria to protect those uses, (3) implemented plans (for needed industrial–municipal waste treatment improvements), and (4) enforcement plans and an anti-degradation statement to protect existing high-quality waters.

REFERENCES

1. Crusberg, Theodore C., et al., "Fungal Biotrap for Retrieval of Heavy Metals from Industrial Wastewater," Third International Conference on New Frontiers for Hazardous Waste Management, Pittsburgh, PA, Sept. 1989.
2. Applications Analysis Report, "Hazcon Solidification Process, Douglasville, Pennsylvania," EPA 540/A5-89/001.
3. Technology Evaluations Report, "SITE Demonstration Test, Hazcon Solidification, Douglasville, Pennsylvania," EPA 540/5-89/001a.
4. Applications Analysis Report, "Shirco Infrared Incineration System," EPA 540/A5-89/010.
5. Technology Evaluation Report, "SITE Program Demonstration Test, Shirco Infrared Incineration System, Peak Oil, Brandon, Florida," EPA/540/5/88-002a.
6. Technology Evaluation Report, "SITE Program Demonstration Test, Shirco Pilot-Scale Infrared Thermal Incineration System, Rose Township Demode Road Superfund Site," EPA 540/5-89/007a.
7. Technology Evaluation Report, "SITE Program Demonstration Test, American Combustion Pyretron Destruction System at the U.S. EPA's Combustion Research Facility," EPA 540/A5-89/008.
8. Applications Analysis Report, "Terra Vac In-Situ Vacuum Extraction System," EPA 540/A5-89/003.
9. Technology Evaluation Report, "SITE Program Demonstration Test, Terra Vac In-Situ Vacuum Extraction System, Groveland Massachusetts," EPA 540/A5-89/003a.
10. Technology Evaluation Report, "SITE Program Demonstration Test, International Waste Technologies In-Situ Stabilization/Solidification, Hialeah, Florida," EPA 540-6-89/004a.
11. "Superfund '89 Conference. November 27–29, 1989," EPA 540/8-89/010.

SUPPLEMENTAL READING

Bachmann, A., et al., "Biodegradation of Alpha- and Beta-Hexachlorocyclohexane in a Soil Slurry under Different Redox Conditions," *Appl. Environ. Microbiol.*, **54**, 143 (1988).

Beall, J. F., and McGathen, R., "Guidelines for Wastewater Treatment I—How to Minimize Wastewater," *Metal Finish.*, **75** (9), 13 (1977); *Metals Abs.*, 11 58–0007 (1978).

Beccari, M., et al., "Kinetics of Dissimilatory NO_3 and NO_2 Reduction in Suspended Growth Culture," *J. Water Pollut. Control Fed.*, **55**, 58 (1983).

Bouwer, E. J., et al., "Anoxic Transformations of Trace Halogenated Aliphatics," in *Toxic and Hazardous Wastes*, G. D. Boardman (Ed.), Proc. 18th Mid-Atl. Ind. Waste Conf., Technomic Publishing Co., Lancaster, PA, 1986.

Bowers, A. R., et al., "Treatment of Toxic Refractory Wastewaters with Hydrogen Peroxide," *Water Sci. Technol.*, **21**, 477 (1989).

Brower, G. R., et al., "Control of Water Problems in a Steel Mill's Blast Furnace Gas Wash Recirculation System," *Proc.* 32nd Ind. Waste Conf., p. 549, Purdue Univ., West Lafayette, IN, 1977.

"Calendar of EPA Regulations under Consideration," *Pollution Eng.*, **10** (12), 61 (1978).

Carins, J., et al., "Developing an On-site Continuous Biological Monitoring System for the Chemical Industry," *Proc.* 5th WWEMA Ind. Poll. Conf., p. 285, WWEMA, McLean, VA, 1978.

Carter, J. L., et al., "Comparison of a Full-Scale Anaerobic Filter with the Pilot Plant Used for Design," *Proc.* Ind. Waste Symp., p. 214, Water Pollut. Control Fed., Alexandria, VA, 1986.

Cole, C. A., et al., "Treatment of Toxic or Refractory Wastewaters with Hydrogen Peroxide," *J. Water Pollut. Control Fed.*, **48**, 297 (1976).

Crawford, G. V., and Teletze, G. H., "Performance of a Hybrid Anaerobic Process," *Proc.* 41st Ind. Waste Conf., p. 196, Purdue Univ., West Lafayette, IN, 1986.

DeAngelis, D. F., et al., "Industrial Waste Pretreatment Using CELROBIC Anaerobic Reactor," *Proc.* 41st Ind. Waste Conf., p. 10, Purdue Univ., West Lafayette, IN, 1986.

"Design Manual for Dewatering Municipal Wastewater Sludges," EPA-625/1-87/014, U.S. EPA, Cincinnati, OH (1987).

Donovan, M. F., et al., "Anaerobic Treatment of Sulfite Liquor Evaporator Condensate," *Proc.* TAPPI Environ. Conf., 1984.

"Economic Indicators," *Chem. Eng.*, **4**, 9 (1988).

Engineering News Record, McGraw-Hill Construct. Weekly, **11**, No. 57 (1988).

Ferguson, J. R., "Impact of EPA Regulations on the Steel Industry," *Pollution Eng.*, **10** (8), 23 (1978).

Gibson, S. A., and Suflita, J. M., "Extrapolation of Biodegradation Results to Groundwater Aquifers: Reductive Dehalogenation of Aromatic Compounds," *Appl. Environ. Microbiol.*, **52**, 681 (1986).

Gray, K. J., "Anaerobic Pharmaceutical Pretreatment System," *Proc.* Second Natl. Conf. Anaerobic Digestion Ind. Wastes, Argonne Natl. Lab., U.S. Dept. Energy, Chicago, IL, Sept. 1986.

Hamza, A., et al., "Potential for Water Reuse in an Egyptian Poultry Processing Plant," *J. Food Sci.*, **43**, 1153 (1978).

"Handbook for Estimating Sludge Management Costs," EPA-625/6-85/010, U.S. EPA, Cincinnati, OH (1985).

Hickman, G. T., and Novak, J. T., "Microcosm Assessment of Biodegradation Rates of Organic Compounds in Soils," in *Toxic and Hazardous Wastes*, J. C. Evans (Ed.), *Proc.* 19 Mid-Atl. Ind. Waste Conf., Technomic Publishing Co., Lancaster, PA (1987).

Howerton, D. E., and Young, J. C., "Two-Stage Cyclic Operation of Anaerobic Filters," *J. Water Pollut. Control Fed.*, **59**, 788 (1987).

Lacy, W. J., "The Closed-Loop Cycle for Industrial Wastewater; The Future Pollution Solution," *Environment International*, **2**, 3–8 (1979).

Mathews, D. G., "Hydrogen Peroxide Controls Odor, Corrosion in Collection Systems," *Water Sew. Works*, **124**, 6 (1977).

McCabe, W. L., et al., *Unit Operations of Chemical Engineering*, McGraw-Hill, New York, 1985.

"Miles Wastewater Pretreatment Plant," Miles Laboratory, Inc., Elkhart, IN (1986).

Miyakoshi, S., "New Aspects of Sludge Incineration in Yokohama," EPA-600/9-84-021, 8th U.S./Jap. Conf. Sewage Treatment Technol., Cincinnati, OH (1984).

Nebel, C., et al., "Ozone Oxidation of Phenolic Effluents," *Proc.* 31st Ind. Waste Conf., Purdue Univ., West Lafayette, IN, 1976; p. 940, Ann Arbor Science Publishers, Inc., Ann Arbor, MI, 1977.

Novak, J. T., et al., "Biodegradation of Methanol and Tertiary Butyl Alcohol in Subsurface Systems," *Water Sci. Technol.*, **17**, 71 (1985).

Oh, Y. M., and Yang, B. S., "Anaerobic Wastewater Treatment Using Floating Media," *Water Sci. Technol.*, **18**, 225 (1986).

Okkes, I. A. G., "Celrobic Process for Anaerobic Treatment of Industrial Wastewater," in *Anaerobic Waste Water Treatment*, W. J. van den Brink (Ed.), p. 299, AWWT Symp. Secretariat, TNO, The Hague, Netherlands (1983).

Ouchi, H., "Study on Fluidized Bed Incineration with a Drying System," EPA-600/9-85-014, 9th U.S./Jap. Conf. Sewage Treatment Technol., Tokyo (1985).

Owen, W. F., et al., "Bioassay for Monitoring Biochemical Methane Potential and Anaerobic Toxicity," *Water Res.*, **13**, 485 (1979).

Ragan, J. L., "Celanese Experience with Anaerobic Filters," in *Anaerobic Filters: An Energy Plus for Wastewater Treatment*, ANL/CNSU-TU-50, p. 129, Argonne Natl. Lab., Argonne, IL, 1980.

Rogers, C. J., "Recycling of Water in Poultry Processing Plants," EPA 600/2-78-039, Cincinnati, OH (1978).

Sheridan, R. P., "The Evolution of Anaerobic Treatment of Industrial Wastewaters," paper presented at the Am. Inst. Chem. Eng. Summer Natl. Meeting, Cleveland, OH, 1984.

Smith, J. A., and Novak, J. T., "Biodegradation of Chlorinated Phenols in Subsurface Aquifers," *Appl. Environ. Microbiol.*, **53**, 710 (1987).

Speece, R. E., "Anaerobic Biotechnology for Industrial Wastewater Treatment," *Environ. Sci. Technol.*, **17**, 416A (1983).

Sulflita, J. M., and Miller, G. D., "Microbial Metabolism of Chlorophenolic Compounds in Ground Water Aquifers," *Environ. Toxicol. Chem.*, **4**, 751 (1985).

Szendrey, L. M., "Startup and Operation of the Bacardi Corporation Anaerobic Filter," *Proc.* Third Int. Symp. Anaerobic Digestion, p. 363 Boston, MA, 1983.

"The Cost Digest: Cost Summaries of Selected Environmental Control Technologies," EPA-600/8-84-010, U.S. EPA, Cincinnati, OH (1984).

Ulrich, G. D., *Guide to Chemical Engineering Process Design and Economics*, John Wiley & Sons, New York, 1984.

van der Berg., L., et al., "Anaerobic Downflow Stationary Fixed-Film Reactors: Performance under Steady-State and Non-Steady-State Conditions," *Proc.* Sem. Anaerobic Fixed-Film Digestion, Pollut. Control Assoc., p. 1, Ontario, Canada, Nov. 1986.

van den Berg, L., and Lentz, C. P., "Effects of Film Area-to-Volume Ratio, Film Support, Height and Direction of Flow on Performance of Methanogenic Fixed-Film reactors," in *Anaerobic Filters: An Energy Plus for Wastewater Treatment*, AHL/CNSU-TU-50, Argonne Natl. Lab., Argonne, IL, 1980.

Vater, W., "Energy Savings Sludge Incineration in Stuttgart's Central Sewage Works, F.R.G.," *Water Sci. Technol.*, **16**, 531 (1984).

Wang, Y. T., et al., "Effects of Preozonation on the Anaerobic Biodegradability of *o*-Cresol," *J. Envir. Eng. Div., Proc. Am. Soc. Civ. Eng.*, **115**, 336 (1989).

Wang, Y. T., et al., "The Effect of Concentration of Phenols on Their Batch Methanogenesis," *Biotechnol. Bioeng.*, **33**, 1353 (1989).

Wilkie, A., et al., "Media Effects in Anaerobic Filters," in *Anaerobic Waste Water Treatment*, W. J. van den Brink (Ed.), p. 242, AWWT Symp. Secretariat, TNO, The Hague, Netherlands (1983).

Williamson, K., and McCarty, P. L., "A Model for Substrate Utilization by Bacterial Films," *J. Water Pollut. Control Fed.*, **48**, 9 (1976).

Williamson, K., and McCarty, P. L., "Verification Studies of the Biofilm Model for Bacterial Substrate Utilization by Bacterial Films," *J. Water Pollut. Control Fed.*, **48**, 9 (1976).

Witt, E. R., et al., "Full-Scale Anaerobic Filter Treats High Strength Waste," *Proc.* 34th Ind. Waste Conf., Purdue Univ., West Lafayette, IN, 1979.

Wittle, T. E., "Design and Startup of an Anaerobic Treatment System for Food Processing Wastewater," paper presented at the Food Process. Waste Conf., Georgia Tech. Res. Inst., Atlanta, GA, 1988.

Woods, D. R., and Anderson, S., "Evaluation of Capital Cost Data: Drives," *Can. J. Chem. Eng.*, **4**, 357 (1975).

Young, H. W., and Young, J. C., "Hydraulic Characteristics of Upflow Anaerobic Filters," *J. Environ. Eng.*, **114**, 621 (1988).

Young, H. W., and Young, J. C., "Full-Scale Treatment of Chemical Process Waste Using Anaerobic Filters," *Proc.* Ind. Waste Symp., Dallas, TX, 1988.

4

Applied Statistical Methods and the Chemical Industry

Robert Kasprzyk* and Stephen Vardeman**

INTRODUCTION

The discipline of statistics is the study of effective methods of data collection, data summarization, and (data based, quantitative) inference making in a framework that explicitly recognizes the reality of nonnegligible variation in many real-world processes and measurements.

The ultimate goal of the field is to provide tools for extracting the maximum amount of useful information about a noisy physical process from a given investment of data collection and analysis resources. It is clear that such a goal is relevant to the practice of industrial chemistry. The primary purposes of this chapter are to indicate in concrete terms the nature of some existing methods of applied statistics that are particularly appropriate to industrial chemistry, and to provide an entry into the large and detailed statistical literature for those readers who find in the discussion here reasons to believe that statistical tools can help them be effective in their work. A secondary purpose is to indicate the relation of applied statistical methodology to modern quality philosophy in the chemical industry.

This chapter will begin with some simple ideas of modern descriptive statistics, including numerical and graphical data summarization tools, and the notions of fitting equations to data and using theoretical distributions. Next, some tools for routine industrial process monitoring and capability assessment, concentrating primarily on the notion of control charting, will be presented. This will be followed by a more extensive discussion of common statistical data collection strategies and data analysis methods for multi-factor experimental situations met in both laboratory and production environments. This section will touch on ideas of partitioning observed variation in a system response to various sources thought to influence the response, factorial and fractional factorial experimental designs, sequential experimental strategy, screening experiments, and response surface fitting and representation. Next come brief discussions of two types of special statistical tools associated specifically with chemical applications, namely, mixture techniques and

*Dow Chemical Company.
**Iowa State University, Department of Statistics.

nonlinear mechanistic model building. A short exposition of chemical industry implications of relationships between modern quality philosophy and the discipline of statistics follows. The chapter concludes with a reference section listing sources for further reading.

SIMPLE TOOLS OF DESCRIPTIVE STATISTICS

Statisticians have developed a variety of data summarization or description methods whose purpose is to make evident the main features of a data set. (Their use, of course, may be independent of whether or not the data collection process actually employed was in any sense a "good" one.) To illustrate some of the simplest of these methods, consider the data listed in Table 4.1. These numbers represent aluminum impurity contents (in ppm) of 26 bihourly samples of recycled PET plastic recovered at a Rutgers University recycling pilot plant.

A simple plot of aluminum content against time order, often called a *run chart*, is a natural place to begin looking for any story carried by a data set. Figure 4.1 shows such a plot for the data of Table 4.1, and in this case reveals only one potentially interesting feature of the data. That is, there is perhaps a weak hint of a downward trend in the aluminum contents that might well have been of interest

TABLE 4.1 Twenty-six Consecutive Aluminum Contents (ppm)[a]

291,	222,	125,	79,	145,	119,	244,	118,	182,	63,	30,	140,	101
102,	87,	183,	60,	191,	119,	511,	120,	172,	70,	30,	90,	115

[a]Based on data in Albin.[1]

Fig. 4.1. A run chart for 26 consecutive aluminum contents.

to the original researchers. (If indeed the possible slight decline in aluminum contents is more than "random scatter," knowledge of its physical origin, whether in actual composition of recycled material or in the measurement process, presumably would have been helpful to the effective running of the recycling facility. We will save a discussion of tools for rationally deciding whether there is more than random scatter in a plot like Fig. 4.1 until the next section.)

The run chart is a simple, explicitly dynamic tool of descriptive statistics. In those cases where one decides that there is in fact little information in the time order corresponding to a data set, there are a variety of simple, essentially static, statistical tools that can be used in describing the pattern of variation in a data set. Figures 4.2 through 4.5 show graphical representations of the data of Table 4.1 in, respectively, *histogram, stem and leaf plot, dot plot,* and *box plot* forms.

The histogram/bar chart idea of Fig. 4.2 is likely familiar to most readers, being readily available, for example, through the use of commercial spreadsheet software. It shows how data are spread out or distributed across the range of values represented, tall bars indicating high frequency or density of data in the interval covered by the base of the bar. Figure 4.2 shows the measured aluminum contents to be somewhat asymmetrically distributed (statistical jargon is that the distribution is "skewed right"), with a "central" value perhaps somewhere in the vicinity of 120 ppm.

Histograms are commonly and effectively used for final data presentation, but as working data analysis tools they suffer from several limitations. In one direction, their

Fig. 4.2. A histogram for 26 aluminum contents.

appearance is fairly sensitive to the data grouping done to make them, and it is usually not possible to recover from a histogram the exact data values used to produce it, should one wish to try other groupings. In another direction, histograms are somewhat unwieldy, for example, not being particularly suitable to the comparison of, say, 10 or 12 data sets on a single page. The graphical devices of Figs. 4.3 through 4.5 are less common than the histogram, but address some of these shortcomings.

```
5 | 11
4 |
4 |
3 |
3 |
2 | 91
2 | 22,44
1 | 72,82,83,91
1 | 01,02,15,18,19,19,20,25,40,45
0 | 60,63,70,79,87,90
0 | 30,30
```

Fig. 4.3. A stem and leaf plot for 26 aluminum contents.

The stem and leaf diagram of Fig. 4.3 and the dot plot of Fig. 4.4 carry shape information about the distribution of aluminum contents in a manner very similar to the histogram of Fig. 4.2. But the stem and leaf and dot diagrams do so without losing the exact identities of the individual data points. The box plot of Fig. 4.5 represents the "middle half" of the data with a box divided at the 50th percentile (or in statistical jargon, the median) of the data, and then uses so-called whiskers to indicate how far the most extreme data points are from the middle half of the data.

Box plots preserve much of the shape information available from the other displays (e.g., portraying lack of symmetry through differing sizes of box "halves" and/or whisker lengths), but do so in a way that is conducive

Fig. 4.4. A dot plot for 26 aluminum contents.

Fig. 4.5. A box plot for aluminum contents.

to simultaneous representation and comparison of many data sets on a single graphic, through the placement of box plots side by side. Figure 4.6 illustrates this point with a graphical comparison of three laboratory test methods to a standard.

A total of 75 samples of a stock solution known to contain 25 parts per million of an impurity were analyzed by a single lab team (25 of the samples being allocated to each of the three methods), and the box plots in Fig. 4.6 portray the measured impurity levels for the different methods. The figure shows quite effectively that Method A is neither precise nor accurate, Method B is quite precise but not accurate, and Method C is somewhat less precise than B but is accurate. This kind of knowledge can form the basis of an informed choice of method.

Figures 4.2 through 4.6 give only a hint of the spectrum of tools of statistical graphics that are potentially helpful in data analysis for industrial chemistry. For more details and much additional reading on the subject of modern statistical graphics, the reader is referred to the book by Chambers et al.[2] listed in the references section.

Complementary to graphical data summaries are *numerical summarizations*. For the simple case of data collected under a single set of conditions, the most commonly used measures deal with the location/center of the data set and the variability/spread of the data. The (*arithmetic*) *mean* and the *median* are the most popular measures of location, and the *variance* and its square root, the *standard deviation*, are the most widely used measures of internal variability in a data set.

For n data values $y_1, y_2, \ldots, y_n$ the median is

$$\tilde{y} = \text{the "middle" or } \frac{n+1}{2}\text{th ordered data value} \quad (4\text{-}1)$$

and the mean is

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \quad (4\text{-}2)$$

The reader is invited to check that upon ordering the $n = 26$ values in Table 4.1, the 13th smallest value is 119 and the 14th smallest value is also 119, so that the only sensible interpretation of (4-1) for the aluminum content data is that

$$\tilde{y} = \text{the 13.5th ordered data value} = \frac{119 + 119}{2} = 119 \text{ ppm}$$

Fig. 4.6. Side-by-side box plots for three laboratory test methods.

On the other hand, from (4-2) the mean of the aluminum contents is

$$\bar{y} = \frac{1}{26}(291 + 222 + 125 + \cdots + 30 + 90 + 115)$$

$$\approx 142.7 \text{ ppm}$$

The media and mean are clearly different measures of location/center. The former is in the middle of the data in the sense that about half of the data are larger and about half are smaller. The latter is a kind of "center of mass," and for asymmetrical data sets like that of Table 4.1 is usually pulled from the median in the direction of any "skew" present, that is, is pulled in the direction of "extreme" values.

The variance of n data values $y_1, y_2, \ldots, y_n$ is essentially a mean squared deviation of the data points from their mean. In precise terms, the variance is

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 \qquad (4\text{-}3)$$

and the so-called standard deviation is

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2} \qquad (4\text{-}4)$$

For the example of the aluminum contents, it is elementary to verify that

$$s^2 \approx \frac{1}{26-1}((291 - 142.7)^2 + (222 - 142.7)^2 + \cdots + (115 - 142.7)^2)$$

$$\approx 9{,}644 \text{ (ppm)}^2$$

so that

$$s = \sqrt{s^2} \approx 98.2 \text{ ppm}$$

An appropriate interpretation of s is not completely obvious at this point, but it does turn out to measure spread of a data set, and to be extremely useful in drawing quantitative inferences from data. (In most, but not all, circumstances met in practice, the *range* or largest value in a data set minus the smallest value is on the order of four to six times s.) The variance and standard deviation are time-honored and fundamental quantifications of the variation present in a single group of measurements and, by implication, the data-generating process that produced them.

When data are collected under several different sets of conditions, and those conditions can be expressed in quantitative terms, effective data summarization often takes the form of *fitting an approximate equation* to the data. As the basis of a simple example of this, consider the data in Table 4.2. The variable x, hydrocarbon liquid hourly space velocity, specifies the conditions under which information on the response variable y, a measure of isobutylene conversion, was obtained in a study involving the direct hydration of olefins.

For purposes of economy of expression, and perhaps some cautious interpolation between values of x not included in the original data set, one might well like to fit a simple equation involving some parameters $\underset{\sim}{b}$, say,

$$y \approx f(x|\underset{\sim}{b}) \qquad (4\text{-}5)$$

to the data of Table 4.2. The simplest possible form for the function $f(x|\underset{\sim}{b})$ that accords with the "up then back down again" nature of the conversion values y in Table 4.2 is the quadratic form

$$f(x|\underset{\sim}{b}) = b_0 + b_1 x + b_2 x^2 \qquad (4\text{-}6)$$

and a convenient method of fitting such an equation (that is linear in the parameters b) is *the method of least squares*. That is, to fit a parabola through a plot of the seven (x, y) pairs specified in Table 4.2 it is convenient to choose b_0, b_1, and b_2 to minimize the sum of squared differences between the observed conversion values y and the corresponding fitted values of y on the parabola. In symbols, the least squares fitting of the approximate

TABLE 4.2 Seven Liquid Hourly Space Velocity/Mole % Conversion Data Pairs[a]

Liquid Hourly Space Velocity, x	*Mole % Isobutylene Conversion*, y
1	23.0, 24.5
2	28.0
4	30.9, 32.0, 33.6
6	20.0

[a]Based on a graph in Odioso *et al.*[3]

relationship specified by (4-5) and (4-6) to the data of Table 4.2 proceeds by minimization of

$$\sum_{i=1}^{7} (y_i - (b_0 + b_1 x_i + b_2 x_i{}^2))^2$$

over choices of the coefficients b. As it turns out, use of standard statistical "regression analysis" software shows that the fitting process for this example produces the approximate relationship

$$y \approx 13.64 + 11.41x - 1.72x^2$$

and Fig. 4.7 shows the fitted (summarizing) parabola sketched on the same set of axes used to plot the seven data points of Table 4.2.

The least squares fitting of approximate functional relationships to data with even multidimensional explanatory variable x typically goes under the (unfortunately obscure) name of *multiple regression* analysis, and is given an introductory treatment in most engineering statistics textbooks, including, for example, the ones by Devore[4] and Vardeman[5] listed in the references. A lucid and rather complete treatment of the subject can also be found in the book by Neter, Wasserman, and Kutner.[6]

A final notion that we wish to treat in this section on descriptive statistics is that of representing a distribution of responses and/or the mechanism that produced them (under a single set of physical conditions) by a *theoretical distribution.* That is, there are a number of convenient theoretical distributional shapes, and it is often possible to achieve great economy of expression and thought by seeing in a graphical representation such as Figs. 4.2 through 4.5 the possibility of henceforth describing the phenomenon portrayed via some one of those theoretical distributions. Here we will concentrate on only the most commonly used theoretical distribution, the so-called *Gaussian* or *normal* distribution.

Figure 4.8 is a graph of the function of x

$$g(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad (4\text{-}7)$$

Fig. 4.7. A scatterplot of seven space velocity/mole % conversion data pairs and a fitted parabola.

Fig. 4.8. The Gaussian probability density with mean μ and standard deviation σ.

where $g(x)$ specifies the archetypical "bell-shaped curve" centered at the number μ, with spread controlled by the number σ (and is in fact usually called the Gaussian probability density with mean μ and standard deviation σ).

Figure 4.8 can be thought of as a kind of idealized histogram. Just as fractional areas enclosed by particular bars of a histogram correspond to fractions of a data set with values in the intervals represented by those bars, areas under the curve specified in (4-7) above particular intervals might be thought of as corresponding to fractions of potential data points having values in those intervals. (It is possible to show that the total area under the curve represented in Fig. 4.8, namely, $\int_{-\infty}^{\infty} g(x)\,dx$, is 1.) Simple tabular methods presented in every elementary statistics book avoid the need to regularly use numerical integration in evaluating such areas. These methods can, for example, be used to show that roughly 68% of a Gaussian distribution lies between $\mu - \sigma$ and $\mu + \sigma$, roughly 95% lies between $\mu - 2\sigma$ and $\mu + 2\sigma$, and roughly 99.7% lies between $\mu - 3\sigma$ and $\mu + 3\sigma$. Part of the convenience provided when one can treat a data-generating process as approximately Gaussian is that, given only a theoretical mean μ and theoretical standard deviation σ, predictions of fractions of future data values likely to fall in intervals of interest are thus easy to obtain.

At this point let us return to the aluminum content data of Table 4.1. The skewed shape that is evident in all of Figs. 4.2 through 4.5 makes a Gaussian distribution inappropriate as a theoretical model for (raw) aluminum content of such PET samples. But as is often the case with right skewed data, considering the *logarithms* of the original measurement creates a scale where a normal distribution is more plausible as a representation of the phenomenon under study. Thus, Table 4.3 contains the natural logs of the values in Table 4.1, and the corresponding stem and leaf plot in Fig. 4.9 shows the transformed data to be much more symmetrically distributed than the original data. The possibility opened up by this kind of transformation idea is one of

TABLE 4.3 Twenty-six Logarithms of Aluminum Contents

5.67,	5.40,	4.83,	4.37,	4.98,	4.78,	5.50,	4.77,	5.20,	4.14,	3.40,	4.94,	4.62
4.62,	4.47,	5.21,	4.09,	5.25,	4.78,	6.24,	4.79,	5.15,	4.25,	3.40,	4.50,	4.74

```
6 | .24
5 | .50,.67
5 | .15,.20,.21,.25,.40
4 | .50,.62,.62,.74,.77,.78,.79,.79,.83,.94,.98
4 | .09,.14,.25,.37,.47
3 |
3 | .40,.40
```

Fig. 4.9. A stem and leaf plot for the logarithms of 26 aluminum contents.

using statistical methods based on the normal distribution to reach conclusions about $\ln y$ and then simply exponentiating to derive conclusions about the original response y itself. The applicability of statistical methods developed for normal distributions is thereby significantly broadened.

In addition to providing convenient conceptual summarizations of the nature of response distributions, theoretical distributions such as the normal distribution form the mathematical underpinnings of methods of formal quantitative *statistical inference*. It is outside our purposes in this chapter to provide a complete introduction to such methods, but thorough and readable accounts are available in engineering statistics books such as those of Devore[4] and Vardeman.[5] Here, we will simply say that, working with a Gaussian description of a response, it is possible to quantify in various ways how much information is carried by data sets of various sizes. For instance, if a normal distribution describes a response variable y, then in a certain well-defined sense, based on $n = 26$ observations producing a mean $\bar{y}$ and a standard deviation s, the interval with end points

$$\bar{y} - 2.060s\sqrt{1 + \frac{1}{26}}$$

and (4-8)

$$\bar{y} + 2.060s\sqrt{1 + \frac{1}{26}}$$

has a 95 percent chance of predicting the value of an additional observation. For instance, applying formula (4-8) to the log values in Table 4.3, the conclusion is that the interval from 3.45 ln(ppm) to 6.10 ln(ppm) has in some sense a 95 percent chance of bracketing an additional log aluminum content produced (under the physical conditions of the original study) at the recycling plant. Exponentiating, the corresponding statement about raw aluminum content is that the interval from 31 ppm to 446 ppm has in some sense a 95 percent chance of bracketing an additional aluminum content. Methods of statistical inference like that represented in (4-8) are called *prediction interval* methods. The book by Hahn and Meeker[7] provides a thorough discussion of such methods, based not only on the Gaussian distribution but on other theoretical distributional shapes as well.

TOOLS OF ROUTINE INDUSTRIAL PROCESS MONITORING AND CAPABILITY ASSESSMENT

Probably the two most basic generic industrial problems commonly approached using statistical methods are those of (1) monitoring and maintaining the stability/consistency of a process and (2) assessing the capability of a stable process. This section provides a brief introduction to the use of tools of "control" charting in these enterprises.

Working at Bell Labs during the 1920s and 1930s, Walter Shewhart developed the notion of routinely plotting data from an industrial process in a form that allows one to separate observed variability in a response into two kinds of variation. The first is that variation which appears to be inherent, unavoidable, short-term, baseline, and characteristic of the process (at least as currently configured). This variation Shewhart called *random* or *common cause variation*. The second kind of variability is that variation which appears to be avoidable, long-term, and/or due to sources outside of those seen as legitimately impacting process behavior. This variation he called *assignable* or *special cause variation*.

Shewhart reasoned that by plotting summary statistics from periodically collected data sets against time order of collection, one would

be able to see interpretable trends or other evidence of assignable variation on the plots, and could intervene to eliminate the physical causes of that variation. The intention was thereby to make process output stable or consistent to within the inherent limits of process precision. As a means of differentiating plotted values that should signal the need for intervention from those that carry no special message of process distress, he suggested drawing so-called control limits on the plots. (The word "control" is something of a misnomer, at least as compared to common modern engineering usage of the word in referring to the active, moment-by-moment steering or regulation of processes. The nonstandard and more passive terminology "monitoring limits" would actually be far more descriptive of the purpose of Shewhart's limits.) These limits were to separate plausible values of the plotted statistic from implausible values when in fact the process was operating optimally, subject only to causes of variation that were part of standard conditions.

By far the most famous implementations of Shewhart's basic logic come where the plotted statistic is either the mean, the range, or, less frequently, the standard deviation. Such charts are commonly known by the names *x-bar charts*, *R charts*, and *s charts*, respectively. As a basis of discussion of Shewhart charts, consider the data given in Table 4.4. These values represent melt index measurements of specimens of extrusion grade polyethylene, taken four per shift in a plastics plant.

TABLE 4.4 Measured Melt Indices for Ten Groups of Four Specimens[a]

Shift	*Melt Index*	$\bar{y}$	R	s
1	218, 224, 220, 231	223.25	13	5.74
2	228, 236, 247, 234	236.25	19	7.93
3	280, 228, 228, 221	239.25	59	27.37
4	210, 249, 241, 246	236.50	39	17.97
5	243, 240, 230, 230	235.75	13	6.75
6	225, 250, 258, 244	244.25	33	14.06
7	240, 238, 240, 243	240.25	5	2.06
8	244, 248, 265, 234	247.75	31	12.92
9	238, 233, 252, 243	241.50	19	8.10
10	228, 238, 220, 230	229.00	18	7.39

[a]Based on data from page 207 of Wadsworth, Stephens, and Godfrey.[8]

Figure 4.10 shows plots of the individual melt indices, means, ranges, and standard deviations from Table 4.4 against shift number. The last three of these are the beginnings of so-called Shewhart $\bar{x}$, R, and s control charts.

What remain to be added to the plots in Fig. 4.10 are appropriate control limits. In order to indicate the kind of thinking that stands behind control limits for Shewhart charts, let us concentrate on the issue of limits for the plot of means. The fact is that mathematical theory suggests how the behavior of *means* $\bar{y}$ ought to be related to the distribution of *individual* melt indices y, provided the data-generating process is stable, that is, subject only to random causes. If individual responses y can be described as normal with some mean μ and standard deviation σ, mathematical theory suggests that averages of n such values will behave as if a different normal distribution were generating them, one with a mean $\mu_{\bar{y}}$ that is numerically equal to μ and with a standard deviation $\sigma_{\bar{y}}$ that is numerically equal to $\sigma/\sqrt{n}$. Figure 4.11 illustrates this theoretical relationship between the behavior of individuals and the behavior of means.

The relevance of Fig. 4.11 to the problem of setting control chart limits on means is that if one is furnished with a description of the typical pattern of variation in y, sensible exceptations for variation in $\bar{y}$ follow from simple normal distribution calculations. So Shewhart reasoned that since about 99.7 percent (most) of a Gaussian distribution is within 3 standard deviations of the center of the distribution, means found to be farther than 3 theoretical standard deviations (of $\bar{y}$) from the theoretical mean (or $\bar{y}$) could be safely attributed to other than chance causes. Hence, furnished with standard values for μ and σ (describing individual observations), sensible control limits for $\bar{y}$ become

$$\text{Upper Control Limit for } \bar{y} = \mu_{\bar{y}} + 3\sigma_{\bar{y}}$$

$$= \mu + 3\frac{\sigma}{\sqrt{n}}$$

Fig. 4.10. Plots of melt index, $\bar{y}$, R, and s against shift number.

Fig. 4.11. The distributions of individuals, y, and sample means, $\bar{y}$.

and

$$\text{Lower Control Limit for } \bar{y} = \mu_{\bar{y}} - 3\sigma_{\bar{y}}$$

$$= \mu - 3\frac{\sigma}{\sqrt{n}} \quad (4\text{-}9)$$

Returning to the context of our example represented by the data of Table 4.4, Wadsworth et al.[8] state that the target value for melt index in the original application was in fact 235. So if standard process behavior is "on target" behavior, the value $\mu = 235$ seems appropriate for use in (4-9). No parallel value for σ was provided by the authors. Common practice in such situations is to use the data in hand (the data of Table 4.4) to produce a plausible value for σ to use in (4-9). There are many possible ways to produce such a value, but to understand the general logic behind the standard ones, it is important to understand what σ is supposed to measure. The variable σ is intended as a theoretical measure of baseline, short-term, common cause variation. As such, the safest way to try to approximate it is to somehow use only measures of variation *within* the groups of four values in Table 4.4 *not* influenced by variation *between* groups. (Measures of variation derived from considering all the data simultaneously, e.g., would reflect variation between shifts as well as the shorter-term variation within shifts.) In fact, the most commonly used ways of obtaining from the data in hand a value of σ for use in (4-9) are based on the averages of the (within-group) ranges or standard deviations. For example, the 10 values of R given in Table 4.4 have a mean

$$\bar{R} = \frac{1}{10}(13 + 19 + 59 + \cdots + 19 + 18) = 24.9$$

and some standard mathematical theory suggests that because the basic group size here is $n = 4$, an appropriate multiple of $\bar{R}$ for use in estimating σ is

$$\frac{\bar{R}}{2.059} \approx 12.1 \quad (4\text{-}10)$$

(The divisor above is a tabled factor commonly called d_2, which increases with n.)

Finally, substituting 235 for μ and 12.1 for σ in (4-9) produces numerical control limits for $\bar{y}$:

$$\text{LCL} = 235 - 3\frac{(12.1)}{\sqrt{4}} = 216.9$$

and

$$\text{UCL} = 235 + 3\frac{(12.1)}{\sqrt{4}} = 253.1$$

Comparison of the $\bar{y}$ values in Table 4.4 to these limits reveals no "out of control" means, that is, no evidence in the means of assignable process variation. Figures 4.12 and 4.13 show control charts for all of $\bar{y}$, R, and s, where control limits for the last two quantities have been derived using standard calculations not shown here.

The R and s charts in Figs. 4.12 and 4.13 are related representations (only one is typically made in practice) of the shift-to-shift behavior of melt index *consistency*. It is seen that on both charts, the shift #3 point plots above the upper control limit. The strong suggestion thus is that melt index consistency was detectably worse on that shift than on the others, so that from this point of view the process was in fact *not stable* over the time period represented in Table 4.4. In practice, physical investigation and hopefully correction of the origin of the instability typically would follow, as well as some reconsideration of our earlier assessment of 12.1 as a plausible figure to represent the inherent short-term variability of melt index. (If shift #3 could be treated as a special case, explainable as an unfortunate but correctable situation that was not expected to reoccur, there might be reason to revise $\bar{R}$ downward by deletion of shift #3 from the calculation, and thereby to reduce one's view of the size of baseline process variability. Notice that, in general, such a downward revision of $\bar{R}$ might well also have the effect of causing one to need to rethink his or her assessment of the constancy of the melt index *mean*.)

There is a variation on the basic "$\bar{x}$ and R chart" idea that we wish to illustrate here next, because of its frequent usefulness in

Fig. 4.12. Control charts for $\bar{y}$ and R based on melt indices.

Fig. 4.13. Control charts for $\bar{y}$ and s based on melt indices.

chemical industry applications. That is the making of a so-called x and MR chart pair. The motivation for this modification of the ideas outlined thus far in this section is that in many chemical process monitoring contexts the natural "group size" is $n = 1$. A mean of $n = 1$ observation(s) is simply that observation itself, and the limits of (4-9) make perfectly good sense for the case of $n < 1$. That is, the analog of an $\bar{x}$ chart for $n = 1$ cases is clear, *at least if one has an externally provided value for* σ. But what, if anything, to do for an $n = 1$ counterpart of the R chart and how to develop an analog of (4-10) in cases where σ is not a priori known are perhaps not so obvious. Table 4.5 contains data representing moisture contents in 0.01 percent of bihourly samples of a polymer, and the question at hand is what besides simply the bihourly y values might be plotted in the style of a Shewhart control chart for such data.

The final column of Table 4.5 gives 19 so-called moving ranges of pairs of successive moisture contents. It is often argued that although these MR values are actually affected not only by variation within a 2-hour production period but by some variation between these periods as well, they come as close to representing purely short-term variation as any measure available from $n = 1$ data. Accordingly, as a kind of $n = 1$ analog of an R chart, moving ranges are often charted in addition to individual values y. Further, the average moving range is used to estimate σ in cases where information on the inherent variability of individuals is a priori lacking, according to the formula

$$\text{estimated } \sigma = \frac{\overline{MR}}{1.128}$$

where $\overline{MR}$ is the mean of the moving ranges (and plays the role of $\bar{R}$ in (4-10)), and 1.128 is the $n = 2$ version of the factor d_2 alluded to immediately below (4-10).

In the case of the data of Table 4.5,

$$\overline{MR} = \frac{1}{19}(16 + 4 + 5 + \cdots + 16 + 18) \approx 8.2$$

so that a (possibly somewhat inflated due to between period variation) data-based estimate of within-period variability σ for use, for example in limits (4-9) is

$$\frac{8.2}{1.128} \approx 7.2$$

Figure 4.14 shows both an x (individuals) chart and an MR (moving range) chart based on these calculations. As no standard value of moisture content was provided in Burr's text,[9] the value $\bar{y} = 28.55$ was used as a substitute for μ in (4-9). The MR chart limits are based on standard $n = 2$ (because ranges of "groups" of two observations are being plotted) r chart control limit formulas. Figure 4.14 shows no evidence of assignable variation in the moisture contents.

The use of Shewhart control charts is admirably documented in a number of statistical quality control books, including those by Wadsworth et al.,[8] Duncan,[10] Burr,[9] Grant and Leavenworth,[11] and Ott and Schilling.[12] Our purpose here is not to provide all details necessary for their use, but only to give the reader an introduction to the overall function that they serve. It should be said, however,

TABLE 4.5 Moisture Contents for 20 Polymer Samples[a]

Sample	*Moisture*, y	*Moving Range*, MR
1	36	—
2	20	16
3	16	4
4	21	5
5	32	11
6	34	2
7	32	2
8	34	2
9	23	11
10	25	2
11	12	13
12	31	19
13	25	6
14	31	6
15	34	3
16	38	4
17	26	12
18	29	3
19	45	16
20	27	18

[a]Based on data from page 190 of Burr.[9]

Fig. 4.14. Control charts for y and MR based on moisture contents of 20 polymer samples.

that in recent years other statistical process monitoring tools such as the so-called CUmulative SUM (CUSUM) schemes and Exponentially Weighted Moving Average (EWMA) schemes have been developed as competing methodologies, and can in some circumstances be practically more effective than the original Shewhart charts. Indeed, many computerized controllers for real-time chemical process monitoring and adjustment now employ some form of CUSUM or EWMA logic. For more on these topics, including their integration with model-based process controllers, the reader is referred to Lucas,[13,14] Lucas and Crosier,[15] Lucas and Saccucci,[16] Crowder,[17,18] Hunter,[19] and, in particular, Vander Wiel et al.[20]

Shewhart's basic conceptualization of common and special cause variation not only leads to control charts as quantitative, rational tools to guide one in knowing when (and when not!) to intervene in an industrial process to correct potential ills, but it also provides a framework for considering the question of what is the best/most consistent performance one can hope for from a particular version of a process. That is, it provides a framework for discussing process capability assessment.

If $\hat{\sigma}$ is some (standard deviation type) estimate of the baseline variation inherent in an industrial process (obtained, e.g., from a calculation such as (4-10) or from data taken from the process after eliminating all physical sources of assignable variation), it essentially specifies what is possible in terms of consistency of process output. There are, however, several common ways of using such an estimate to produce related measures of process capability.

For one thing, remembering again the fact that an interval from $\mu - 3\sigma$ to $\mu + 3\sigma$ (i.e., of length 6σ) will bracket about 99.7 percent of a normal distribution, the figure $6\hat{\sigma}$ is sometimes stated as "the process capability." This usage would say that in the context of the polyethylene melt index example of Table 4.4 the $\hat{\sigma} = 12.1$ figure from (4-10) implies a melt index process capability of $6 \cdot (12.1) \approx$

72.6. If properly monitored, the process appears capable of producing almost all individual melt indices in a 73-point range.

Where there are stated specifications for individual measurements y, $\hat{\sigma}$ is sometimes turned into a kind of index comparing it to the difference in upper and lower engineering specifications. For example, one such *process capability index* is

$$C_p = \frac{\text{USL} - \text{LSL}}{6\hat{\sigma}}$$

where USL − LSL is the difference in specifications. Fairly obviously, the larger the value of C_p, the more comfortably (properly targeted) process output values will fit in an interval from LSL to USL.

Another process capability measure that is frequently used in the industrial chemistry sector is

$$C_{pk} = \text{minimum}\left\{C_{pu} = \frac{\text{USL} - \hat{\mu}}{3\hat{\sigma}},\ C_{pl} = \frac{\hat{\mu} - \text{LSL}}{3\hat{\sigma}}\right\}$$

where $\hat{\mu}$ is an estimated overall process average for an in-control/stable/predictable process, and $\hat{\sigma}$ is as before. This measure is clearly similar to C_p, but it takes into account the placement of the process mean in a way that is ignored by C_p. A large value of C_{pk} indicates that not only is the process short-term variation small enough for the process output values to potentially fit comfortably between LSL and USL, but that the process is currently so targeted that the potential is being realized.

STATISTICAL METHODS AND INDUSTRIAL EXPERIMENTATION

Probably the most important area of opportunity for the new application of statistical methods in the chemical industry during the 1990s is that of increasing the effectiveness of industrial experimentation. That is, it is one thing to bring an existing industrial process to stability (a state of "statical control"), but it is quite another to determine how to make fundamental changes in that process that will improve its basic behavior. This second activity almost always involves some form of experimentation, whether it be in the laboratory or in a plant. As we indicated in the introduction, efficient methods and strategies of such data collection (and corresponding analysis) are a central concern of applied statistics. In this section, we hope to give the reader some insight into the kinds of statistical tools that are available for use in chemical industry experimentation.

We will here take as our meaning of the term "experimentation" the observation of a (typically noisy) physical process under more than one condition, with the broad goal of understanding and then using knowledge of how the process reacts to the changes in conditions. In most industrial contexts, the "conditions" under which the process is observed can be specified in terms of the settings or so-called levels chosen for several potentially important process or environmental variables, the so-called factors in the experiment. In some cases, the hope is to identify those (often largely unregulated) factors and combinations of factors that seem to most influence an observed response variable, as a means of targeting them for attention intended to keep them constant or otherwise to eliminate their influence, and thereby to improve the consistency of the response. In other situations the hope is to discover patterns in how one or more critical responses depend on the levels of (often tightly controlled) factors, in order to provide a road map for the advantageous guiding of process behavior (e.g., to an increased mean reaction yield) through enlightened changing of those levels.

This section is organized into two subsections. In the first, we will illustrate the notion of variance component estimation through an example of a nested or hierarchical data collection scheme. In the second, we will discuss some general considerations in the planning of experiments to detail the pattern of influence of factors on responses, consider so-called factorial and fractional factorial

experimental designs, illustrate response surface fitting and interpretation tools and the data requirements they imply, and, in the process, discuss the integration of a number of statistical tools in a sequential learning strategy.

Identifying Major Contributors to Process Variation

A statistical methodology that is particularly relevant where experimentation is meant to identify important unregulated sources of variation in a response is that of variance component estimation, based on so-called ANalysis Of VAriance (ANOVA) calculations and random effects models. As an example of what is possible, consider the data of Table 4.6. Shown here are copper content measurements for some bronze castings. Two copper content determinations were made on each of two physical specimens cut from each of 11 different castings.

The data of Table 4.6 were *by design* collected to have a "tree type" or so-called hierarchical/nested structure. Figure 4.15 shows a diagram of a generic hierarchical structure for *balanced* cases like the present one, where there are equal numbers of branches leaving all nodes at a given level (there are equal numbers of determinations for each specimen and equal numbers of specimens for each casting).

An important goal in most hierarchical studies is determining the size of the contributions to response variation provided by the different factors, that is, the different levels of the tree structure. (In the present context, the issue is how variation between castings compares to variation between specimens within a casting, and how they both compare to variation between determinations for a given specimen. If the overall variability observed were considered excessive, such analysis could then help guide efforts at variation reduction by identifying the largest contributors to observed variability.) The structure portrayed in Fig. 4.15 turns out to enable an appealing statistical analysis, providing help in that quantification.

TABLE 4.6 Forty-four Copper Content Measurements From 11 Bronze Castings[a]

Casting	*Specimen*	*Determination*	*Copper Content*, y (%)
1	1	1	85.54
1	1	2	85.56
1	2	1	85.51
1	2	2	85.54
2	1	1	85.54
2	1	2	85.60
2	2	1	85.25
2	2	2	85.25
3	1	1	85.72
3	1	2	85.77
3	2	1	84.94
3	2	2	84.95
4	1	1	85.48
4	1	2	85.50
4	2	1	84.98
4	2	2	85.02
5	1	1	85.54
5	1	2	85.57
5	2	1	85.84
5	2	2	85.84
6	1	1	85.72
6	1	2	85.86
6	2	1	85.81
6	2	2	85.91
7	1	1	85.72
7	1	2	85.76
7	2	1	85.81
7	2	2	85.84
8	1	1	86.12
8	1	2	86.12
8	2	1	86.12
8	2	2	86.20
9	1	1	85.47
9	1	2	85.49
9	2	1	85.75
9	2	2	85.77
10	1	1	84.98
10	1	2	85.10
10	2	1	85.90
10	2	2	85.90
11	1	1	85.12
11	1	2	85.17
11	2	1	85.18
11	2	2	85.24

[a]Based on data taken from Wernimont.[21]

Fig. 4.15. A balanced hierarchical data structure.

If one lets

y_{ijk} = the copper content from the kth determination of the jth specimen from casting i

$\bar{y}_{ij\cdot} = \frac{1}{2}\sum_k y_{ijk}$ = the mean copper content determination from the jth specimen from casting i

$\bar{y}_{i\cdot\cdot} = \frac{1}{2}\sum_j \bar{y}_{ij\cdot}$ = the mean copper content determination from the ith casting

and

$\bar{y}_{\cdots} = \frac{1}{11}\sum_i \bar{y}_{i\cdot\cdot}$ = the overall mean copper determination

it is possible to essentially break down the variance of all 44 copper contents (treated as a single group) into interpretable pieces, identifiable as variation between $\bar{y}_{i\cdot\cdot}$'s (casting means), variation between $\bar{y}_{ijk\cdot}$'s (specimen means) within castings, and variation between y_{ijk}'s (individual measurements) within a specimen. That is, it is an algebraic identity that for 44 numbers y_{ijk} with the same structure as those in Table 4.6

$$(44-1)s^2 = \sum_{i,j,k}(y_{ijk} - \bar{y}_{\cdots})^2$$

$$= \sum_{i,j,k}(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdots})^2 + \sum_{i,j,k}(\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot})^2$$

$$+ \sum_{i,j,k}(y_{ijk} - \bar{y}_{ij\cdot})^2 \qquad (4\text{-}11)$$

The sums indicated in (4-11) are over all data points; so, for example, the first summand on the right is obtained for the copper content data by summing each $(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdots})^2$ a total of $2 \cdot 2 = 4$ times, one for each determination on a given casting. With the obvious meaning for the $\bar{y}$'s and the substitution of the total number of data values for 44, the identity in (4-11) applies to any balanced hierarchical data structure. It is a so-called ANOVA identity, providing an intuitively appealing partitioning of the overall observed variability in the data, an *analyzing of the (observed) variation.*

Some tedious arithmetic "by hand," or by use of nearly any commercially available statistical package that includes an ANOVA program, shows that for the copper content data of Table 4.6 the numerical version of (4-11) is approximately

$$5.1385 = 3.2031 + 1.9003 + .0351 \qquad (4\text{-}12)$$

Although we will not provide any details here, the reader is alerted to the fact that it is common practice to present the elements of an identity such as (4-12) in a tabular form called an "ANOVA table." The use for the elements of (4-12) that we wish to illustrate here is their role in estimating casting, specimen, and determination "variance components."

That is, if one models an observed copper determination as the sum of a *random* casting-dependent *effect* whose distribution is described by a variance σ_c^2, a *random* specimen-dependent *effect* whose distribution is described by a variance σ_s^2, and a *random*

determination-dependent *effect* whose distribution is described by a variance σ_d^2, the elements of (4-12) lead to estimates of the *variance components* σ_c^2, σ_s^2, and σ_d^2 in the model. Note that in such a random effects model of the data-generating process, copper measurements from the same casting share the same casting effect, and copper measurements from the same specimen share the both same casting and the same specimen effects. The individual σ^2 values are conceptually the variances that would be seen in copper contents if only the corresponding sources of variation were present. The sum of the σ^2 values is conceptually the variance that would be seen in copper contents if single determinations were made on a number of different castings.

Standard statistical methodology for estimation of the variance components (which we will not detail here, but can, e.g., be found in Chapter 11 of Hicks[22] produces

$$\hat{\sigma}_d^2 = \frac{.0351}{11 \cdot 2 \cdot (2-1)} \approx .0016\,(\%)^2$$

as an estimate of σ_d^2,

$$\hat{\sigma}_s^2 = \frac{1}{2}\left(\frac{1.9003}{11 \cdot (2-1)} - .0016\right) \approx .0856\,(\%)^2$$

as an estimate of σ_s^2, and

$$\hat{\sigma}_c^2 = \frac{1}{2 \cdot 2}\left(\frac{3.2031}{(11-1)} - \frac{1.9003}{11 \cdot (2-1)}\right) \approx .0369\,(\%)^2$$

as an estimate of σ_c^2. Figure 4.16 is a pie chart representation of these three estimated variance components as fractions of their sum (the variance predicted if single determinations were made on single specimens from each casting), and graphically identifies inhomogeneity between specimens cut from a single casting as the biggest contributor to observed variation.

On the standard deviation scale the estimates translate to $\hat{\sigma}_d \approx 0.04\%$, $\hat{\sigma}_d \approx 0.29\%$, and $\hat{\sigma}_c \approx 0.19\%$. So, for example, the data of Table 4.6 indicate that even if castings and specimens were all exactly alike, it would still be reasonable to expect measured copper contents to vary according to a standard deviation of about .04 percent, presumably due to unavoidable measurement error.

Fig. 4.16. Three estimated variance components for copper contents.

Variance component estimation methodology is not limited to balanced hierarchical experiments, but they do provide an important and straightforward context in which to introduce the technology. More detailed information on the case discussed here and extensions to other kinds of data structures can be found in books by Vardeman,[5] Neter et al.,[6] Mason, Gunst, and Hess,[23] and Hicks.[22]

Discovering and Exploiting Patterns of Factor Influence on Responses

Having discussed statistical methodology particularly appropriate to studies whose primary purpose is simply to identify factors with the largest influence on a response, we will now consider methods aimed more directly at detailed experimental quantification of the pattern of factor influence on one or more responses. As an example, we will use a "sanitized" account of some statistical aspects of a highly successful and economically important process improvement project. (Data presented here are not the original data, but resemble them in structure. Naturally, details of the project not central to our expository purposes and those of a proprietary nature will be suppressed.)

The process monitoring, capability assessment, and variance source identification ideas discussed thus far are almost logical prerequisites for industrial experimentation to detail the nature of dependence of response variables on factors of interest. When an industrial process has been made to operate in a stable manner, its intrinsic variability reduced to the extent practically possible, and that baseline performance quantified and understood, the prospects of success are greatly enhanced in subsequent efforts to understand the effects of potential fundamental process changes.

Preliminary work by various groups left a project team with a batch production process behaving in a stable but unsatisfactory fashion. Obvious sources of variation (both in the process itself and "upstream") has been identified and, to the degree practically possible, eliminated. The result was a process with an average output *purity* of 88 percent and an associated purity standard deviation of around 5 percent, and an average *yield* of 43 percent and an associated yield standard deviation of around 5 percent as well. The project team was charged with finding ways to increase the purity and yield means to, respectively, 95 percent and 59 percent while hopefully also further reducing the standard deviations. To accomplish this, the team recognized the need for an improved understanding of how various process variables under their control influenced purity (which we will call y_1) and yield (which we will call y_2). Experimentation to provide this was authorized, and, in particular, attention was focused on four factors consisting of three reactant concentrations and the process run time. We will call the Reactant A mole ratio x_1, the Reactant B mole ratio x_2, the Reactant C mole ratio x_3, and the run time (in hours) x_4.

The choice of experimental factors (what to vary in data collection) is a nontrivial matter of fundamental importance that is best handled by people with firsthand process knowledge. There are a number of popular techniques and tools (such as so-called cause and effect diagrams, discussed for instance in the book by Wadsworth et al.[8]) for helping groups brainstorm and reach a consensus on such matters. Further, in cases where a priori knowledge of a process is scarce, relatively small preliminary screening experiments can help reduce a large list of potential factors to a smaller list apparently worthy of more detailed study. (The fractional factorial plans that will be illustrated shortly often are recommended for this purpose.)

Once a particular set of experimental factors has been identified, questions about exactly how they should be varied must be answered. To begin with, there is the choice of levels for the factors, the matter of how much the experimental factors should be varied. Particular experimental circumstances usually dictate how this is addressed. Widely spaced (substantially different) levels will in general lead to bigger changes in responses, and therefore clearer indications of how the responses depend upon the experimental factors, than will closely spaced (marginally different) levels. But they may do so at the expense of potentially creating unacceptable or even disastrous process conditions or output. Thus, what may be an acceptable strategy in a laboratory study might be completely unacceptable in a production environment and vice versa.

Given a set or range of levels for each of the individual experimental factors, there is still the question of exactly what combinations of levels actually will be used to produce experimental data. For example, in the process improvement study, standard process operating conditions were $x_1 = 1.5$, $x_2 = 1.15$, $x_3 = 1.75$, and $x_4 = 3.5$, and the project team decided on the ranges

$$1.0 \leqslant x_1 \leqslant 2.5,\ 1.0 \leqslant x_2 \leqslant 1.8,\ 1.0 \leqslant x_3 \leqslant 2.5, \text{ and } \quad (4\text{-}13)$$

$$2.0 \leqslant x_4 \leqslant 5.0$$

as defining the initial limits of experimentation. But the question remained as to exactly what sets of mole ratios and corresponding run times were appropriate for data collection.

A natural, but largely discredited, strategy of data collection is the one-variable-at-a-time

experimental strategy of picking some base of experimental operations (such as standard operating conditions) and varying the level of only one of the factors away from that base at a time. The problem with such a strategy is that sometimes two or more factors act on responses jointly, doing things in concert that neither will do alone. For example, in the process improvement study, it might well have been that an increase in either x_1 or x_2 alone would have affected yield very little, whereas a simultaneous increase in both would have caused an important increase. Modern strategies of industrial experimentation are conceived with such possibilities in mind, and attempt to spread out observations in a way that gives one some ability to identify the nature of the response structure no matter how simple or complicated it turns out to be.

There are several issues to consider when planning the combinations of levels to include in an experiment. We have already said that it is important to "vary several factors simultaneously." It also is important to provide for some replication of at least a combination or two in the experiment, as a means of getting a handle on the size of the experimental error or baseline variation that one is facing. The replication both verifies the reproducibility of values obtained in the study and identifies the limits of that reproducibility. Also, one must balance the urge to "cover the waterfront" with a wide variety of combinations of factor levels against resource constraints and a very real law of diminishing practical returns as one goes beyond what is really needed in the way of data to characterize response behavior. In addition, the fact that real-world learning is almost always of a sequential rather than a "one shot" nature suggests that it is in general wise to spend only part of an experimental budget on early study phases, leaving resources adequate to follow up directions suggested by what is learned in those stages.

It is obvious that a minimum of two different levels of an experimental factor must appear in a set of experimental combinations if any information is to be gained on the effects of that factor. So one logical place to begin thinking about a candidate design for an industrial experiment is with the set of all possible combinations of two levels of each of the experimental factors. If there are p experimental factors, statistical jargon for such an arrangement is to call it a (complete) $2 \times 2 \times \cdots \times 2$ or 2^p factorial plan. For example, in the process improvement situation, an experiment consisting of the running of all 16 possible combinations of

$$x_1 = 1.0 \text{ or } x_1 = 2.5$$

$$x_2 = 1.0 \text{ or } x_2 = 1.8$$

$$x_3 = 1.0 \text{ or } x_3 = 2.5$$

and

$$x_4 = 2.0 \text{ or } x_4 = 5.0$$

would be called a complete $2 \times 2 \times 2 \times 2$ or 2^4 factorial experiment. Notice that in geometric terms, the (x_1, x_2, x_3, x_4) points making up this 2^4 structure amount to the 16 "corners" in four-dimensional space of the initial experimental region defined in (4-13).

A complete factorial experimental plan is just that, in some sense "complete." It provides enough information to allow one to assess (for the particular levels used) not only individual but also joint or interaction effects of the factors on the response or responses. But when in fact (unbeknownst to the investigator) a system under study is a relatively simple one, principally driven by only a few individual or low-order joint effects of the factors, fewer data actually are needed to characterize those effects adequately. So what is often done in modern practice is initially to run only a carefully chosen part of a full 2^p factorial, a so-called fractional factorial plan, and to decide based on the initial data whether data from the rest of the full factorial appear to be needed in order adequately to characterize and understand response behavior. We will not here discuss the details of how so-called 2^{p-q} fractional factorials are intelligently chosen, but there is accessible reading material on the subject in books by Box, Hunter, and Hunter,[24] and by Vardeman.[5]

In the process improvement study, what was actually done in the first stage of data collection was to gather information from one-half of a full 2^4 factorial (a 2^{4-1} fractional factorial) augmented by four observations at the "center" of the experimental region (thereby providing both some coverage of the interior of the region, in addition to a view of some of its corners, and important replication as well). The data in Table 4.7 are representative of what the group obtained.

The order in which the data are listed is simply a convenient systematic one, not to be confused with the order in which experimental runs were actually made. The table order is far too regular for it to constitute a wise choice itself. For example, the fact that all $x_3 = 1.0$ combinations precede the $x_3 = 2.5$ ones might have the unfortunate effect of allowing the impact of unnoticed environmental changes over the study period to end up being confused with the impact of x_3 changes. The order in which the 12 experimental runs were actually made was chosen in a "completely randomized" fashion. For a readable short discussion of the role of randomization in industrial experimentation, the reader is referred to Box.[25]

For purposes of this discussion, attention is focused on the yield response variable, y_2. Notice first that the four y_2 values from the center point of the experimental region has $\bar{y} = 53.2$ and $s = 1.13$ (which incidentally already appear to be an improvement over typical process behavior). As a partial indication of the logic that can be used to investigate whether the dependence of yield on the experimental factors is simple enough to be described adequately by the data of Table 4.7, one can compute some estimated "main effects" from the first eight data points. That is, considering first the impact of the variable x_1 (alone) on yield, the quantity

$$\begin{aligned}\bar{y}_{\text{high}\,x_1} - \bar{y}_{\text{low}\,x_1} &= \tfrac{1}{4}(45.9 + 46.0 + 51.2 + 52.8) \\ &\quad - \tfrac{1}{4}(35.1 + 4.0 + 41.4 + 10.0) \\ &= 26.35\end{aligned}$$

is perhaps a sensible measure of how a change in x_1 from 1.00 to 2.50 is reflected in yield. Similar measures for the other variables turn out to be

$$\bar{y}_{\text{high}\,x_2} - \bar{y}_{\text{low}\,x_2} = -15.20$$

$$\bar{y}_{\text{high}\,x_3} - \bar{y}_{\text{low}\,x_3} = 6.10$$

and

$$\bar{y}_{\text{high}\,x_4} - \bar{y}_{\text{low}\,x_4} = 0.45$$

These measures provide some crude insight into the directions and magnitudes of influence of the experimental variables on y_2. (Clearly, by these measures $x_1 = 2.50$ seems preferable to $x_1 = 1.00$, and the run time variable x_4 seems to have little impact on yield.) But they also provide strong evidence that the nature of the dependence of yield on the experimental factors is too complicated to be described by the action of the factors individually. For example, if it *were* the case that the separate actions of the experimental factors were adequate to describe system behavior, then standard statistical theory and the data indicate that the mean response for the $x_1 = 1.00$, $x_2 = 1.0$, $x_3 = 1.00$, and $x_4 = 2.0$ set of conditions would be around

$$\begin{aligned}\hat{y} = \bar{y}_{\text{corners}} &- \tfrac{1}{2}(26.35) - \tfrac{1}{2}(-15.20) \\ &- \tfrac{1}{2}(6.10) - \tfrac{1}{2}(0.45) = 27.45\end{aligned}$$

(where $\bar{y}_{\text{corners}}$ is standing for the mean of the first eight yields in Table 4.7). But the observed yield of 35.1 is clearly incompatible

TABLE 4.7 Data from an Initial Phase of a Process Improvement Study

x_1	x_2	x_3	x_4	*Purity,* y_1*(%)*	*Yield,* y_2*(%)*
1.00	1.0	1.00	2.0	62.1	35.1
2.50	1.0	1.00	5.0	92.2	45.9
1.00	1.8	1.00	5.0	7.0	4.0
2.50	1.8	1.00	2.0	84.0	46.0
1.00	1.0	2.50	5.0	61.1	41.4
2.50	1.0	2.50	2.0	91.6	51.2
1.00	1.8	2.50	2.0	9.0	10.0
2.50	1.8	2.50	5.0	83.7	52.8
1.75	1.4	1.75	3.5	87.7	54.7
1.75	1.4	1.75	3.5	89.8	52.8
1.75	1.4	1.75	3.5	86.5	53.3
1.75	1.4	1.75	3.5	87.3	52.0

with such a mean and the standard deviation value (of $s = 1.13$) derived from the repeated center point. Also, other simple evidence that (at least linear and) separate action of the four factors is not enough to describe yield adequately is given by the large difference between $\bar{y}_{\text{corners}} = 35.8$ and the observed mean from the center point $\bar{y} = 53.2$. (As it turns out, calculations that we will not show here indicate the *possibility* that individual action of the factors *plus joint action* of the Reactant A and Reactant B mole ratios is sufficient to describe yield. But in any case, the point is that the data of Table 4.7 provide evidence that the pattern of dependence of yield on the experimental variables is not simple, and thus that completion of the 2^4 factorial is in order.)

After a complete analysis of the first round of experimental data, the project team "ran the second half fraction" of the 2^4 factorial, and data similar to those in Table 4.8 were obtained. (Again, no significance should be attached to the order in which the observations in Table 4.8 are listed. It is not the order in which the experimental runs were made.)

The data from the second phase of experimentation served to complete the project team's 2^4 factorial picture of yield and confirm the tentative understanding drawn first from the initial half fraction. It is seen that the combinations listed in Table 4.8 are in the same order as the first eight in Table 4.7 as regards levels of experimental variables x_1, x_2, and x_3, and that the corresponding responses are very similar. (This, by the way, has the happy practical implication that run time seems to have little effect on final purity or yield, opening the possibility of reducing or at least not increasing the standard run time.) Thorough data analysis of a type not shown here left the project team with a clear (and quantified version of the) understanding that Reactant A and B mole ratios have important individual and joint effects on the responses, and that, acting independently of the other two reactants, Reactant C also has an important effect on the responses. However, it did *not* yet provide a solution to the team's basic problem, which was to reach a 59 percent mean yield goal.

TABLE 4.8 Data from a Second Phase of a Process Improvement Study

x_1	x_2	x_3	x_4	*Purity,* y_1*(%)*	*Yield,* y_2*(%)*
1.00	1.0	1.00	5.0	64.0	35.3
2.50	1.0	1.00	2.0	91.9	47.2
1.00	1.8	1.00	2.0	6.5	3.9
2.50	1.8	1.00	5.0	86.4	45.9
1.00	1.0	2.50	2.0	63.9	39.5
2.50	1.0	2.50	5.0	93.1	51.6
1.00	1.8	2.50	5.0	6.8	9.2
2.50	1.8	2.50	2.0	84.6	54.3

The data of Table 4.7 and 4.8 do hold out hope that conditions producing the desired purity and yield can be found. That is, though none of the 16 corners of the experimental region nor the center point appeared to meet the team's yield goal, the data do show that there is substantial *curvature* in the yield response. (The joint effect of x_1 and x_2 amounts to a kind of curvature, and the nonlinearity of response indicated by a large difference between $\bar{y}_{\text{corners}} \approx 35.8$ and $\bar{y} = 53.2$ at the center of the experimental region also is a kind of curvature.) If one could "map" the nature of the curvature, there is at least the possibility of finding favorable future operating conditions in the interior of the initial experimental region defined in (4-13).

It ought to be at least plausible to the reader that 2^4 factorial data (even supplemented with center points) are not really sufficient to interpolate the nature of a curved response over the experimental region. More data are needed, and a standard way of augmenting a 2^p design with center points to one sufficient to do the job is through the addition of so-called star points to produce a central composite design. Star points are points outside the original experimental region whose levels of all but one of the p experimental factors match those of the center point. Figure 4.17 shows graphical representations of central composite designs in $p = 2$ and $p = 3$ factors.

The project team conducted a third phase of experimentation by adding eight star points to their study and obtained data similar to those in Table 4.9.

Fig. 4.17. $p = 2$ and $p = 3$ central composite designs.

The data in Tables 4.7, 4.8, and 4.9 taken together turn out to provide enough information to enable one rather thoroughly to quantify the "curved" nature of the dependence of y_2 on x_1, x_2, x_3, and x_4. A convenient and often successful method of accomplishing this quantification is through the least squares fitting of a general *quadratic response surface.* That is, central composite data are sufficient to allow one to fit an equation to a response that involves a constant term, linear terms in all the experimental variables, quadratic terms in all of the experimental variables, and cross-product terms in all pairs of the experimental variables. Appropriate use of a multiple regression program with the project data represented here produces the fitted equation

$$\begin{aligned} y_2 \approx\ & 15.4 + 37.9x_1 - 66.2x_2 + 48.8x_3 \\ & + 0.97x_4 - 16.1x_1^2 - .03x_2^2 \\ & - 13.6x_3^2 - .046x_4^2 + 26.5x_1x_2 \\ & + 0.344x_1x_3 - .217x_1x_4 \\ & + 1.31x_2x_3 - .365x_2x_4 + .061x_3x_4 \end{aligned}$$

This may not seem to the reader to be a particularly helpful data summary, but standard multiple regression tools can be used to deduce that an essentially equivalent, far less cluttered, and more clearly interpretable representation of the relationship is:

$$\begin{aligned} y_2 \approx\ & 13.8 + 37.8x_1 - 65.3x_2 + 51.6x_3 \\ & - 16.2x_1^2 - 13.6x_3^2 + 26.5x_1x_2 \qquad (4\text{-}14) \end{aligned}$$

Equation (4-14), provides an admirable fit to the data in Tables 4.7, 4.8, and 4.9, is in perfect agreement with all that has been said thus far about the pattern of dependence of yield on the experimental factors, *and* allows one to do some intelligent interpolation in the initial experimental region. Use of an equation like (4-14) ultimately allowed the project team to determine that an increase of x_1 only would, with minimal change in the existing process, allow them to meet their yield goal. (In fact, the single change in x_1 proved to be adequate to allow them to meet all of their yield *and* purity goals!)

Graphical representations similar to those in Figs. 4.18 and 4.19 for (4-14) with $x_3 = 1.75$

TABLE 4.9 Data from a Third Phase of a Process Improvement Study

x_1	x_2	x_3	x_4	*Purity*, y_1(%)	*Yield*, y_2(%)
0.6895	1.4	1.75	3.5	20.8	13.0
2.8105	1.4	1.75	3.5	95.9	54.3
1.75	0.8344	1.75	3.5	99.9	62.4
1.75	1.9656	1.75	3.5	65.9	41.2
1.75	1.4	0.6895	3.5	64.4	32.7
1.75	1.4	2.8105	3.5	64.8	40.3
1.75	1.4	1.75	1.379	88.1	52.7
1.75	1.4	1.75	5.621	88.9	50.5

Fig. 4.18. A contour plot of fitted yield when $x_3 = 1.75$.

Fig. 4.19. A perspective graph of fitted yield when $x_3 = 1.75$.

(the standard operating value for x_3) were instrumental in helping the team understand the message carried by their data and how yield could be improved. Figure 4.18 is a so-called contour plot (essentially a topographic map) of the fitted equation, and Fig. 4.19 is a more three-dimensional–looking representation of the same surface. Both types of display are commonly used tools of modern statistical experiment design and analysis. The contour plot idea is particularly helpful where several responses are involved, and by overlaying several such plots one can simultaneously picture the various implications of a contemplated choice of process conditions.

SPECIAL STATISTICAL TOOLS FOR CHEMICAL APPLICATIONS

The statistical methods discussed thus far are of a quite general nature, routinely finding application beyond the bounds of the chemical industry. In this section, we will briefly highlight two statistical methodologies whose most important applications are to chemical problems. That is, we will touch on some of the ideas of mixture experiments and the role of statistics in mechanistic modeling.

Mixture Experiments

In many situations in industrial chemistry, some important measured property of a product is a function of the proportions in which a set of p ingredients or components is represented in a mixture leading to the product. For example, case studies in the literature have covered subjects ranging from octanes of gasoline blends, discussed by Snee;[26] to strengths of different formulations of ABS pipe compound, treated in Koons and Wilt;[27] to aftertaste intensities of different blends of artificial sweeteners used in an athletic sport drink, discussed by Cornell;[28] to moduli of elasticity of different rocket propellant formulations, considered by Kurotori.[29] For experimenting in such contexts, special statistical techniques are needed. These tools have been discussed at length by Cornell,[30,31] and our purpose here is not to attempt a complete exposition, but only to whet the reader's appetite for further reading in this area.

The goal of mixture experimentation is to quantify how proportions $x_1, x_2, x_3, \ldots, x_p$ of ingredients 1 through p affect a response y. Usually, the hope is to fit some kind of approximate equation involving some parameters $\underset{\sim}{b}$, say

$$y \approx f(x_1, x_2, \ldots, x_p | \underset{\sim}{b})$$

to a set of n $(x_1, x_2, \ldots, x_p, y)$ data points, for the purpose of using the fitted equation to guide optimization of y, that is, to find the "best" blend. The logic of data collection and equation fitting is complicated in the mixture scenario by the fact that

$$x_1 + x_2 + \cdots + x_p = 1 \qquad (4\text{-}15)$$

The linear constraint (4-15) means that (p way) factorial experimentation is impossible, and that special measures must be employed in order to use standard regression analysis software to do least squares equation fitting. We will briefly describe in turn some approaches to experimental design, equation fitting, and presentation of results for the mixture problem under its fundamental constraint (4-15).

In the case of $p = 3$ (a three-component mixture problem), the set of all possible combinations of values for x_1, x_2, and x_3 satisfying (4-15) can be conveniently represented as an equilateral triangular region. Figure 4.20 shows such a region and the so-called simplex coordinate system on the region. The corners on the plot stand for cases where the "mixture" involved is actually a single pure component. Points on the line segments bounding the figure represent two-component mixtures, and interior points represent genuine three-component mixtures. For example, the center of the simplex corresponds to a set of conditions where each component makes up exactly one-third of the mixture.

One standard mixture (experimental) design strategy is to collect data at the extremes (corners) of the experimental region along with collecting data on a regular grid in the

Fig. 4.20. The set of points with $x_1 + x_2 + x_3 = 1$ and a simplex coordinate system.

experimental region. Figure 4.21 shows a $p = 3$ example of such a so-called simplex lattice design, and Table 4.10 lists the (x_1, x_2, x_3) points involved. (As in the cases of the data in tables 4.7, 4.8, and 4.9, the order used in the listing in Table 4.10 is not one that would be used in sequencing data collection runs. Instead, a randomly chosen order often is employed.)

Another standard mixture experiment strategy is the so-called simplex centroid design, where data are collected at the extremes of the experimental region and for every equal-parts two-component mixture, every equal-parts three-component mixture,

Fig. 4.21. A $p = 3$ simplex lattice.

TABLE 4.10 (x_1, x_2, x_3) Points in a Particular $p = 3$ Simplex Lattice Design

x_1	x_2	x_3
1	0	0
0	1	0
0	0	1
$\frac{1}{3}$	$\frac{2}{3}$	0
$\frac{2}{3}$	$\frac{1}{3}$	0
$\frac{1}{3}$	0	$\frac{2}{3}$
$\frac{2}{3}$	0	$\frac{1}{3}$
0	$\frac{1}{3}$	$\frac{2}{3}$
0	$\frac{2}{3}$	$\frac{1}{3}$
$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

and so on. Figure 4.22 identifies the blends included in a $p = 3$ simplex centroid design.

Often, the space of practically feasible mixtures is smaller than the entire set of $x_1, x_2, \ldots, x_p$ satisfying (4-15). (For example, in many contexts, "pure" mixtures do not produce viable product. Concrete made using only water and no sand or cement obviously is a useless building product. One common type of constraint on the proportions $x_1, x_2, \ldots, x_p$ that produces quite simple experimental regions is that of lower bounds on one or more of the individual proportions. Cornell,[30] for example, discusses a situation where the effectiveness in grease stain removal of a $p = 3$ bleach mixture was studied; and past experience with the product indicated that the proportions by weight of bromine, x_1, of powder, x_2, and of HCl, x_3, needed to satisfy the following constraints:

$$x_1 \geqslant .30,\ x_2 \geqslant .25, \text{ and } x_3 \geqslant .02 \quad (4\text{-}16)$$

for effective action of the product (i.e., the mixture needed to be at least 30 percent bromine, at least 25 percent powder, and at least 2 percent HCl by weight.)

The effect of adding the lower bound constraints (4-16) to the basic mixture constraint (4-15) can be pictured as in Fig. 4.23. There, a triangular subregion of the basic $p = 3$ simplex depicts the feasible (x_1, x_2, x_3) points. The choice of experimental mixtures for such an experimental region can be made by direct analogy to or rescaling of designs such as the simplex lattice and simplex centroid designs illustrated above to cover the entire simplex. (It is common to refer to the rescaling process as the use of pseudo-components.)

Fig. 4.23. The $p = 3$ simplex and a set of feasible bleach mixtures.

Constraint systems more complicated than simple lower bounds produce irregularly shaped experimental regions and less obvious methods of choosing $(x_1, x_2, \ldots, x_p)$ points to cover the experimental region. When $p = 3$, it is possible to sketch the region of feasible points on a simplex plot and use it to help guide the choice of mixture experiment strategy. Figure 4.24 illustrates the kind of region that can arise with other than exclusively lower bound constraints.

When more than three components are involved in a mixture study, such plots are, of course, no longer possible, and other more analytic methods of identifying candidate experimental mixtures have been developed. For example, McLean and Anderson[32] presented an algorithm for locating the vertices of an experimental region defined by the basic constraint (4-15) and any combination of upper and or lower bound constraints

$$0 \leqslant a_i \leqslant x_i \leqslant b_i \leqslant 1$$

Fig. 4.22. A $p = 3$ simplex centroid design.

Fig. 4.24. An irregularly shaped experimental region in a $p = 3$ mixture study.

on the proportions x_i. Cornell[30,31] discusses a variety of algorithms for choosing good mixture experiment designs under constraints, and many of the existing algorithms for the problem have been implemented in the MIXSOFT software package developed by Piepel.[33]

Empirical polynomial descriptions of (approximately) how a response y depends upon proportions $x_1, x_2, \ldots, x_p$ are popular mixture analysis tools. The process of fitting polynomials to mixture experiment data in principle uses the same least squares notion illustrated in the fitting of a parabola to the data of Table 4.2. However, the mechanics of using standard multiple regression analysis software in the mixture context is complicated somewhat by the basic constraint (4-15). For example, in view of (4-15) the basic ($p + 1$ parameter) linear relationship

$$y \approx b_0 + b_1x_1 + b_2x_2 + \cdots + b_px_p \quad (4\text{-}17)$$

is in some sense "overparameterized" in the mixture context, in that it is equivalent to the (p parameter) relationship

$$y \approx b_1x_1 + b_2x_2 + \cdots + b_px_p \quad (4\text{-}18)$$

if one identifies the coefficients in (4-18) with the sums of the corresponding coefficients in (4-17) and the coefficient b_0. As a result, it is the "no intercept" relationship (4-18) that is typically fit to mixture data when a linear relationship is used. In a similar way, when a second-order or (multivariable) quadratic relationship between the individual proportions and the response variable is used, it has no intercept term and no pure quadratic terms. For example, in the $p = 3$ component mixture case, the general quadratic relationship typically fit to mixture data is

$$y \approx b_1x_1 + b_2x_2 + b_3x_3 + b_4x_1x_2 + b_5x_1x_3 + b_6x_2x_3 \quad (4\text{-}19)$$

(Any apparently more general relationship involving an intercept term and pure quadratic terms can by use of (4-15) be shown to be equivalent to (4-19) in the mixture context.) Relationships of the type of (4-19) are often called Scheffé models, after the first author to treat them in the statistical literature. Other more complicated equation forms are also useful in some applications, but we will not present them in this chapter. The interested reader is again referred to Cornell[30,31] for more information on forms that have been found to be tractable and effective.

We should point out that the ability to fit equations of the form (4-18) or like (4-19), or of an even more complicated form, is predicated on having data from enough different mixtures to allow unambiguous identification of the parameters $\underset{\sim}{b}$. This requires proper data collection strategy. Much of the existing statistical research on the topic of mixture experiment design has to do with the question of wise allocation of experimental resources under the assumption that a particular type of equation is to be fit.

One's understanding of fitted polynomial (and other) relationships often is enhanced through the use of contour plots made on coordinate systems like that in Fig. 4.25. (This is even true for $p \geqslant 3$ component mixture scenarios, but the use of the idea is most transparent in the three-component case.) A plot like Fig. 4.25 can be a powerful tool to aid one in understanding the nature of a fitted equation and finding regions of optimum fitted response.

Fig. 4.25. A contour plot made on the $p = 3$ simplex.

The mixture experiment counterpart to conventional screening/fractional factorial experimentation also is possible. So-called axial designs have been developed for the purpose of providing screening-type mixture data for use in rough evaluation of the relative effects of a large number of mixture components on a response variable. The same kind of sequential experimental strategy illustrated in the process improvement example is applicable in mixture contexts, as well as contexts free of a constraint such as (4-15).

Mechanistic Model Building

The kinds of equations most easily fit to multi-factor data using standard (least squares) regression analysis techniques are polynomial equations such as (4-6), (4-14), (4-18), and (4-19). These are particularly convenient because they are linear in their parameters, b. But they are probably best thought of as empirical "mathematical French curve" descriptions of the relation of a response, y, to the explanatory variables, x. Polynomials equations function as extremely useful summaries of observed patterns in one's data, but they do not typically provide direct insight into chemical mechanisms that produce those patterns, and the fitted parameters, b, do not often have direct physical meanings. Their use is particularly appropriate where there is little a priori knowledge of mechanisms involved in a process that might aid in its description, and/or no such knowledge is really essential to achieving one's goals.

Sometimes, however, it is desirable (on the basis of possible reaction kinetics or for other reasons) to posit theoretical descriptions of process outputs in terms of explanatory variables. That is, physicochemical principles often lead (through differential or integral equation descriptions of a system) to equation forms for a response that, like the following

$$y = \frac{K_1 K_A K_B P_A P_B}{(1 + K_A P_A + K_B P_B)^2}$$

$$y = C_0 \exp(-Kt)$$

and

$$y = \frac{K_1 x}{1 + K_2 x}$$

are nonlinear in the parameters. Although such equations or models may be less tractable than empirical polynomial equations, the parameters involved more often than not *do have* direct physical interpretations. Further, when such a model can be verified as being an adequate description of a process (thereby confirming scientific understanding) and the parameters involved are estimated from process data, such mechanistic models can provide much safer extrapolations beyond an experimental region than the cruder empirical polynomial models.

The process of research in chemical systems is one of developing and testing different models for process behavior. Whether empirical or mechanistic models are involved, the discipline of statistics provides data-based tools for discrimination between competing possible models, parameter estimation, and model verification for use in this enterprise. In the case where empirical models are used, techniques associated with "linear" regression (linear least squares) are used, whereas in mechanistic modeling contexts "nonlinear" regression (nonlinear least squares) techniques most often are needed. In either case, the statistical tools are applied most fruitfully in iterative strategies such as that outlined in Fig. 4.26.

Reilly and Blau[34] and Chapter 16 of Box et al.[24] provide introductions to the general philosophy of using statistical methods in

Fig. 4.26. An iterative model-building strategy. (*From "Simusolv, a Computer Program for Building Mathematical Models," G. Blau and K. Kuenker,* CAST Communications, ***12***, *1, 10–21 (1989). By permission.*)

mechanistic modeling contexts, as well as a number of useful references for further reading.

Fairly sophisticated and specialized statistical software is needed in the practical application of nonlinear regression methods to mechanistic modeling for industrial chemistry applications. The techniques implemented in such software are best discussed in Bates and Watts,[35] Bard,[36] and Reilly and Blau.[34]

MODERN QUALITY PHILOSOPHY AND THE DISCIPLINE OF STATISTICS

It is widely recognized that the well-being of modern business concerns, both within and outside the chemical sector, is closely related to corporate attitudes toward quality. The writing and corporate seminars of respected experts such as Deming,[37] Juran,[38,39] Crosby,[40] and Feigenbaum[41] all promote a modern business paradigm based on the

importance of product and process quality and emphasizing the importance of continuous improvement. These emphases are applied not only to the creation of new products and delivery of existing products, but to all internal and external aspects of business and interaction with suppliers and customers. Such a *total quality* emphasis gives organizations the ability to compete in a global marketplace, and its proponents all point to the important role of management's leadership in the integration of quality and continuous improvement emphases into corporate life.

The discipline of statistics plays an important role in the total quality approach. Perhaps most fundamentally, the conceptual framework that it provides for recognizing and quantifying variation (itself almost a synonym for lack of quality), for using data to give insight into the operation of processes, and for experimentally probing for ways of improving processes is essential to the success of total quality management. Also, the regular enlightened use of the tools of statistical methods of data collection and analysis is the surest known path to the efficient implementation of total quality goals.

The whole gamut of statistical methods, from the simple to the highly sophisticated, is potentially applicable to the implementation of total quality objectives. See, for example, the books of Juran and Gryna,[42] Duncan,[10] Montgomery,[43] Feigenbaum,[41] and Scherkenbach[44] in this regard. The simplest methods, of course, have seen the most widely used. Ishikawa[45] and Kane[46] present some of the simplest of statistically related quality improvement tools, which have been widely embraced and have appeared as key components of many corporate quality improvement success stories. In addition to the run charts, histograms, and control charts illustrated earlier in this chapter, Ishikawa helped popularize the use of *Pareto charts*, *cause and effect* (or *fishbone*) *diagrams*, *flow charts*, and *scatterplots* in quality improvement projects.

Figure 4.27 illustrates a Pareto chart, which is a kind of bar graph intended to help distinguish the "important few" causes of quality problems from the "trivial many."

Cause and effect diagrams such as Fig. 4.28 are helpful tools in organizing a discussion of possible causes of waste and variation.

Flow charts such as Fig. 4.29 are another tool for looking at the steps involved in a work flow, and for stimulating and organizing a discussion of possible causes of variation and waste.

Fig. 4.27. A Pareto chart.

Fig. 4.28. A cause and effect diagram.

Fig. 4.29. A flow chart.

Fig. 4.30. Two scatterplots.

Scatterplots such as Fig. 4.30 provide an initial graphical look at the relationship between two variables, and can provide suggestions of possible patterns of causality (either between the two plotted variables or from a third one, potentially influencing both of the ones plotted).

It should be emphasized that although statistical methods like these simple tools may traditionally be most strongly associated with manufacturing applications, they have in recent years been shown to be highly effective in a variety of service and managerial decision-making applications as well. See, for example, Nolan and Provost[47] and Wheeler and Chambers.[48]

The adoption of statistically based corporate total-quality thinking and tools is, of course, easier said than done. The goal must be embraced by the entire organization. To be effective, the approach must be a part of all employees' daily work process. Such widespread cooperation in a total quality endeavor requires more than simple attendance at a training class on statistical methods. The total quality philosophy must be promoted at all levels of the organization, and proper use of appropriate tools must be guided and nurtured. The books of Scholtes,[49] Hradeskey,[50] Scherkenbach,[44] and Feigenbaum[41] discuss issues and approaches in the implementation of total quality programs.

CONCLUSION

We have tried in this chapter to give the reader the flavor of modern applied statistical methods and to illustrate their usefulness in the chemical industry. Details of their implementation have of necessity been reserved for further, more specialized reading, for which the interested reader is encouraged to consult the references given in this chapter. The bibliographies in Vardeman and Cornell,[51] Vardeman and David,[52] and Hahn and Meeker[53] provide an effective additional source of references for the topics discussed here in general terms.

REFERENCES

1. Albin, S., "The Lognormal Distribution for Modeling Quality Data When the Mean Is Near Zero," *Journal of Quality Technology*, **22**, 105–110 (1990).
2. Chambers, J., Cleveland, W., Kleiner, B., and Tukey, P., *Graphical Methods of Data Analysis*, Duxbury, Boston, MA, 1983.
3. Odioso, R., Henke, A., Stauffer, H., and Frech, K., "Direct Hydration of Olefins," *Industrial and Engineering Chemistry*, **53** (3), 209–211 (1961).
4. Devore, J., *Probability and Statistics for Engineering and the Sciences* (3rd ed.), Brooks/Cole, Pacific Grove, CA, 1991.
5. Vardeman, S., *Statistics for Engineering Problem Solving*, PWS-Kent, Boston, MA, to appear (1993).
6. Neter, J., Wasserman, W., and Kutner, M., *Applied Linear Statistical Models* (3rd ed.), Irwin, Homewood, IL, 1990.
7. Hahn, G., and Meeker, W., *Statistical Intervals: A Guide for Practitioners*, Wiley, New York, 1991.
8. Wadsworth, H., Stephens, K., and Godfrey, B., *Modern Statistical Methods for Quality Control and Improvement*, Wiley, New York, 1986.
9. Burr, I., *Statistical Quality Control Methods*, Dekker, New York, 1976.
10. Duncan, A., *Quality Control and Industrial Statistics* (5th ed.), Irwin, Homewood, IL, 1986.
11. Grant, E., and Leavenworth, R., *Statistical Quality Control* (6th ed.), McGraw-Hill, New York, 1988.

12. Ott, E., and Schilling, E., *Process Quality Control*, McGraw-Hill, New York, 1990.
13. Lucas, James, "The Design and Use of V-Mask Control Schemes," *Journal of Quality Technology*, **8**, 1–8 (1976).
14. Lucas, J., "Combined Shewhart-CUSUM Quality Control Schemes," *Technometrics*, **24**, 51–59 (1982).
15. Lucas, J., and Crosier, R., "Fast Initial Response for CUSUM Quality Control Schemes," *Technometrics*, **24**, 199–205 (1982).
16. Lucas, J., and Saccucci, M., "Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements," *Technometrics*, **32**, 1–12 (1990).
17. Crowder, S., "A Simple Method for Studying Run Length Distributions of Exponentially Weighted Moving Average Charts," *Technometrics*, **29**, 401–407 (1987).
18. Crowder, S., "Design of Exponentially Weighted Moving Averrage Schemes," *Journal of Quality Technology*, **21**, 155–162 (1989).
19. Hunter, J. Stuart, "The Exponentially Weighted Moving Average," *Journal of Quality Technology*, **18**, 203–210 (1986).
20. Vander Wiel, S., Tucker, W., Faltin, F., and Doganaksoy, N., "Algorithmic Statistical Process Control: Concepts and an Application," *Technometrics*, to appear (1992).
21. Wernimont, G., "Statistical Quality Control in the Chemical Laboratory," *Quality Engineering*, **2**, 59–72 (1989).
22. Hicks, C., *Fundamental Concepts in the Design of Experiments* (3rd ed.), Holt, Rinehart and Winston, New York, 1982.
23. Mason, R., Gunst, R., and Hess, J., *Statistical Design and Analysis of Experiments*, Wiley, New York, 1989.
24. Box, G., Hunter, W., and Hunter, J. S., *Statistics for Experimenters*, Wiley, New York, 1978.
25. Box, G., "George's Column," *Quality Engineering*, **2**, 497–502 (1990).
26. Snee, R., "Developing Blending Models for Gasoline and Other Mixtures," *Technometrics*, **23**, 119–130 (1981).
27. Koons, G., and Wilt, M., "Design and Analysis of an ABS Pipe Compound Experiment," in *Experiments in Industry: Design, Analysis and Interpretation of Results*, Ronald Snee, Lynne Hare, and J. Richard Trout (Eds.), pp. 111–117, ASQC Quality Press, Milwaukee, WI, 1985.
28. Cornell, J., "A Comparison between Two Ten-Point Designs for Studying Three-Component Mixture Systems," *Journal of Quality Technology*, **18**, 1–15 (1986).
29. Kurotori, I., "Experiments with Mixtures of Components Having Lower Bounds," *Industrial Quality Control*, **22**, 592–596 (1966).
30. Cornell, J., *How to Run Mixture Experiments for Product Equality*, Vol. 5 in the ASCQ "Basic References in Quality Control" series, American Society for Quality Control, Milwaukee, WI, 1983.
31. Cornell, J., *Experiments with Mixtures: Designs, Models, and the Analysis of Mixture Data* (2nd ed.), Wiley, New York, 1990.
32. MacLean, R., and Anderson, V., "Extreme Vertices Design of Mixture Experiments," *Technometrics*, **8**, 447–454 (1966).
33. Piepel, G., *Mixsoft and Misoft User's Guide, Version 1.0*, Mixsoft—Mixture Experiment Software, Richland, WA, 1989.
34. Reilly, P., and Blau, G., "The Use of Statistical Methods to Build Mathematical Models of Chemical Reacting Systems," *Canadian Journal of Chemical Engineering*, **52**, 289–299 (1974).
35. Bates, D., and Watts, D., *Nonlinear Regression Analysis and Its Applications*, Wiley, New York, 1988.
36. Bard, Yonathan, *Nonlinear Parameter Estimation*, Academic Press, New York, 1974.
37. Deming, W. E., *Out of the Crisis*, MIT Center for Advanced Engineering Study, Cambridge, MA, 1986.
38. Juran, J., *Juran on Planning for Quality*, Free Press, New York, 1988.
39. Juran, J., *Juran on Leadership for Quality: An Executive Handbook*, Free Press, New York, 1989.
40. Crosby, P., *Quality Is Free: The Art of Making Quality Certain*, McGraw-Hill, New York, 1979.
41. Feigenbaum, A., *Total Quality Control* (3rd ed.), McGraw-Hill, New York, 1983.
42. Juran, J., and Gryna, F. (Eds.), *Quality Control Handbook* (4th ed.), McGraw-Hill, New York, 1988.
43. Montgomery, D., *Statistical Quality Control*, Wiley, New York, 1991.
44. Scherkenbach, W., *The Deming Route to Quality and Productivity: Roadmaps and Roadblocks*, ASQC Quality Press, Milwaukee, WI, 1986.
45. Ishikawa, K., *A Guide to Quality Control* (2nd ed.), Asian Productivity Association, Tokyo, 1986.
46. Kane, V., *Defect Prevention: Use of Simple Statistical Tools*, Dekker, New York, 1989.
47. Nolan, T., and Provost, L., "Understanding Variation," in *Continuous Improvement of Quality: An Overview*, Associates in Process Improvement, Houston, TX, 1989.
48. Wheeler, D., and Chambers, D., *Understanding Statistical Process Control*, Statistical Process Controls, Knoxville, TN, 1986.
49. Scholtes, P., (Ed.), *The Team Handbook: How to Use Teams to Improve Quality*, Joiner Associates, Madison, WI, 1988.
50. Hradesky, J., *Productivity and Quality Improvement: A Practical Guide to Implementing Statistical Process Control*, McGraw-Hill, New York, 1988.

51. Vardeman, S., and Cornell, J., "A Partial Inventory of Statistical Literature on Quality and Productivity through 1985," *Journal of Quality Technology*, **19**, 90–97 (1987).
52. Vardeman, S., and David, H., "Statistics for Quality and Productivity: A New Graduate Level Statistics Course," *The American Statistician*, **38**, 225–243 (1984).
53. Hahn, G., and Meeker, W., "An Engineer's Guide to Books on Statistics and Data Analysis," *Journal of Quality Technology*, **16**, 196–218 (1984).

5

Safety Considerations in the Chemical Process Industries

Stanley M. Englund*

INTRODUCTION

In recent years there has been an increased emphasis on chemical process safety as a result of a number of serious accidents. Public awareness of these accidents has provided a driving force for industry to improve its safety record. Local and national governments are taking a hard look at safety in industry as a whole and the chemical industry in particular, and there has been an increasing amount of government regulation.

Although the chemical industry is one of the safest industries, its safety image in the eyes of the public is dismal, perhaps because sometimes when there is an accident in a chemical plant, it is spectacular and receives a great deal of attention. For many reasons, the public often associates the chemical industry with environmental problems and safety problems; and, unfortunately, all too often the negative image that goes with the problems is deserved.

Hazards from combustion and runaway reactions play a leading role in many chemical process accidents. Knowledge of these reactions is essential for the control of process hazards. Much of the damage and loss of life in chemical accidents results from a sudden release of material at high pressures, which may or may not result from fire; so it is important to understand how sudden pressure releases can occur. They can be due, for example, to ruptured high pressure tanks, runaway reactions, flammable vapor clouds, or pressure developed from external fire. Fires can cause severe damage from thermal radiation. Chemical releases from fires and pressure releases can form toxic clouds that can be dangerous to people over large areas. Static electricity often is a hidden actor in accidents. It is important to understand the reactive nature of the chemicals involved in a chemical facility.

*The Dow Chemical Co., Midland, MI.

Chemical process safety involves both the technical and the management aspects of the chemical industry, and this chapter addresses both. It is not enough to be aware of how to predict the effect of process hazards and how to design systems to reduce the risks of these hazards. It also is important to consider how chemical process safety can be managed. Technical people at all levels have a degree of management responsibility, and can contribute to the overall management of safe plants.

Loss of containment due to mechanical failure or misoperation is not a major cause of chemical process accidents. The publication *One Hundred Largest Losses: A Thirty Year Review of Property Damage Losses in the Hydrocarbon-Chemical Industry*[1] cites loss of containment as the leading cause of property loss in the chemical process industries.

The term, "inherent" means "belonging by nature, or the essential character of something." An "inherently safe" plant is safe by its nature and by the way it is constituted.

No facilities can be completely "inherently safe," but they can be made "inherently safer" by careful examination of all aspects of plant design and management. It is possible to identify "inherently unsafe" equipment and management practices and try to avoid them. It has been found that the basic or root cause of most chemical process accidents can be traced to some failure of a management system. Human error plays an important role in many process accidents and is often closely related to problems with management systems. Major efforts are under way to address the potential problems of human error. By being aware of modern advances in design and management techniques, and putting these techniques into practice, it is possible to have facilities that are safer than ever.

The term "intrinsic" has a meaning similar to that of "inherent," but the common usage of "intrinsic" in the chemical industry usually means a protection technique related to electricity. Intrinsic safety is based on the restriction of electrical energy to a level below which sparking or heating effects cannot ignite an explosive atmosphere (Lees,[2] p. 513).

CONSIDERATIONS IN THE DESIGN AND OPERATION OF INHERENTLY SAFER PLANTS*

The design of chemical plants to be more nearly "inherently safe" has received a great deal of attention in recent years. This is due in part to the worldwide attention to safety issues in the chemical industry brought on by the gas release at the Union Carbide plant in Bhopal, India, in December 1984. This and the fairly frequent occurrence of other chemical plant incidents has raised the issue of chemical plant safety to a very high level of visibility and concern. The major factors which should be considered in the planning, design, and operation of chemical plants are described below. The reader is referred to the list of recommended supplementary reading at the end of the chapter for more detailed information.

Responsibility for Safety in Design and Operation

It is very important that responsibility for the safe design and operation of a plant be clearly defined in the early stages. This means that competent and experienced people should be made responsible and held accountable for decisions made from the start of plant design on through plant start-up and operation.

Review of Design Alternatives

Hazards should be considered and eliminated in the process development stage where possible. This would include considerations of alternative processes, reduction or elimination of hazardous chemicals, site selection, etc. By the time the process is developed, the process designers already have major constraints imposed on them.

Hazards should also be identified and removed or reduced early in the design.

*Portions of this section have been reprinted by permission of the publisher from "Opportunities in the Design of Inherently Safer Chemical Plants" by Stanley M. Englund in *Advances in Chemical Engineering*, Volume 15. Copyright 1990 by Academic Press, Inc. (reference 3 in this chapter).

Adding protective equipment at the end of the design or after the plant is operating can be expensive and not entirely satisfactory.

Allowing time in the early stages of design for critical reviews and evaluation of alternatives would involve studies such as an early hazard and operability (HAZOP) study, using flowsheets, before final design begins.[4] Fault tree analysis," QRA (quantitative risk assessment), checklists, audits, and other review and checking techniques can also be very helpful. These techniques are extensively discussed in the technical literature and will not be discussed in detail here.

Emergency Planning

Emergency planning is primarily for the protection of plant personnel and people in nearby areas and the environment that could be affected by plant problems. It should be considered early in the design and should be coordinated with the existing site emergency plan.

Emergency planning includes tornado and storm shelters, flood protection, earthquakes, proximity to public areas, and safe exit routes. It also includes the effect that an emergency in the "new process" would have on other plants, and the effect that an emergency in another plant would have on the new process. The effects of potential spills on waterways and aquifers should be considered.

Emergency response planning is discussed fully in Chapter 6.

Placement of Process and Storage Areas

The Bhopal plant of Union Carbide was built originally 1.5 miles from the nearest housing (see below under "Case Histories"). Over time, a residential area grew up next to the plant. This demonstrates the need to prevent hazardous plants from being located close to residential areas and to prevent residential areas from being established near such plants.[4] If possible, the cost of a plant should include an adequate buffer zone unless other means can be provided to ensure that the public will not build adjacent to the plant. The nature and size of this buffer zone depend on many factors, including the amount and type of chemicals stored and used.

Storage of Hazardous Materials

The best way to minimize releases of a hazardous or flammable material is to have less of it around. In the Flixborough disaster[2] on June 1, 1974, the process involved the oxidation of cyclohexane to cyclohexanone by air (with added nitrogen) in the presence of a catalyst (see below under "Case Histories"). The cyclohexanone was converted to caprolactam, which is the basic raw material for Nylon 6. The reaction product from the final reactor contained approximately 94 percent unreacted cyclohexane at 155°C and over 200 psig. The holdup in the reactors was about 240,000 lb, of which about 80,000 lb escaped. It is estimated that about 20,000 to 60,000 lb actually was involved in the explosion. The resulting large unconfined vapor cloud explosion (or explosions—there may have been two) and fire killed 28 people and injured 36 at the plant and many more in the surrounding area, demolished a large chemical plant, and damaged 1821 houses and 167 shops. The very large amount of flammable liquid well above its boiling and flash points contributed greatly to the extreme severity of the disaster (Lees,[2] p. 863).

The results of the Flixborough investigation make it clear that the large inventory of flammable material in the process plant contributed to the scale of the disaster. It is concluded that "limitations of inventory (of flammable materials) should be taken as specific design objectives in major hazard installations." It should be noted that reduction of inventory may require more frequent and smaller shipments, and this would entail more chances for errors in connecting and reconnecting. These possibly "negative" benefits should also be analyzed. Quantitative risk analysis of storage facilities has revealed solutions that may run counter to intuition.[5]

For example, contrary to popular opinion, reducing inventories in tanks of hazardous materials does little to reduce risk in situations where most of the exposure arises from the number and extent of valves, nozzles, and lines connecting the tank. Removing tanks from the service altogether, on the other hand, helps.

A large tank may offer greater safety than several small tanks of the same aggregate capacity because there are fewer associated nozzles and lines. Also, a large tank is inherently more robust, or can economically be made more robust by deliberate overdesign, than is a small tank and therefore is more resistant to external damage. On the other hand, if the larger tank has larger connecting lines, the relative risk may be greater if release rates through the larger lines increase the risk more than the tank's inherently greater strength reduces it. In transportation of hazardous materials, the benefits of head shields and shelf couplers have been shown through quantitative risk assessment. Maintaining tank car integrity in a derailment is often the most important line of defense in transportation of hazardous materials.

Liquefied Gas Storage

Usually, leaks of liquefied gases are much less serious if such gases are stored refrigerated at low temperatures and pressures than if they are stored at ambient temperatures under pressure. A leak of a volatile liquid held at atmospheric temperature and pressure results in only a relatively slow evaporation of the liquid. Escape of a refrigerated liquefied gas at atmospheric pressure gives some initial flashoff, and then it evaporates at a rate which is relatively slow but faster than the first case, depending on weather conditions.

Loss of containment of a liquefied gas under pressure and at atmospheric temperature, however, causes immediate flashing of a large proportion of the gas, followed by slower evaporation of the residue. This is usually a more serious case than if a refrigerated tank is used. The hazard from a gas under pressure is normally much less in terms of the amount of material stored, but the physical energy released if a confined explosion occurs at high pressure is large.

Use of Open Structures

There are many examples of serious fires and explosions that probably resulted in part from handling moderate to large quantities of flammable or combustible liquids and liquefied flammable gases inside of enclosed structures. If a sufficient quantity of flammable mixture should ignite inside an ordinary chemical processing building, it is highly probable an explosion will occur that will seriously damage the building. For this reason, processing equipment is often installed in a structure without walls, usually called an "open structure." This permits effective ventilation by normal wind currents and aids the dispersion of any vapors that do escape. If ignition of gas occurs within the structure, the absence of walls minimizes the pressure developed from the combustion and the probability of flying shrapnel from a shattered structure.[6]

Substantial explosion damage will be done to a building by combustion of a surprisingly small quantity of a flammable gas–air mixture. If there is an explosion of a flammable gas–air mixture in a building where the flammable gas mixture occupies a space equal to only one or two percent of the building volume, the building may be seriously damaged if it does not have adequate explosion venting. This is because most buildings will suffer substantial structural damage from an internal pressure appreciably less than 1 psi (0.07 bar). Thus, a building does not need to be "full" or even close to "full" of a flammable mixture for a building explosion to occur that can cause considerable damage.

In 1950 a serious hydrocarbon explosion occurred in an enclosed chemical processing unit of The Dow Chemical Company. It was instrumental in causing Dow to establish a policy of using open structures for chemical processes which use substantial quantities of flammable liquids and liquefied flammable gases, and combustible liquids above their flash points.

Need to Understand Reactive Chemicals Systems

The main business of most chemical companies is to manufacture products through the control of reactive chemicals. The reactivity that makes chemicals useful can also make them hazardous. Therefore, it is essential that a process designer understand the nature of the reactive chemicals involved in his or her process.

Usually reactions are carried out without mishaps, but sometimes chemical reactions get out of control for many reasons, including:

1. Wrong raw material used.
2. Operating conditions changed.
3. Time delays unanticipated.
4. Equipment failure.
5. Reactants contaminated.
6. Materials of construction wrong.
7. Misoperation.
8. External fire.

Such mishaps can be worse if the chemistry is not fully understood. A chemical plant can be "inherently safer" only if knowledge of the reactive chemicals systems is carefully used in its design.

Reactive Hazard Evaluations. Reactive hazard evaluations should be made on all new processes, and on all existing processes on a periodic basis. There is no substitute for experience, good judgment, and good data in evaluating potential hazards. Reviews should include:

- Process chemistry:
  Reactions
  Potential pressure buildup
  Intermediate streams
  Side reactions
  Heat of reaction
- Reactive chemicals test data for evidence of:
  Flammability characteristics
  Shock sensitivity
  Exotherms
  Other evidence of instability
- Planned operation of process, especially:
  Upsets
  Delays
  Critical instruments and controls
  Mode of failure
  Redundancy
  Worst credible case scenarios

Worst-Case Thinking. At every point in the operation, the process designer should conceive of the *worst* possible combination of circumstances that could *realistically* exist, such as:

Air leakage
Deadheaded pumps
Instrument failure
Loss of agitation
Loss of compressed air
Loss of cooling water
Plugged lines
Power failure
Raw material impurities
Wrong combination or amount of reactants
Wrong valve position

An engineering evaluation should then be made of the worst-case consequences, with the goal that the plant will be safe even if the worst case occurs. A HAZOP study could be used to help accomplish "worst-case thinking." When the process designers know what the worst-case conditions are, they should:

- Try to avoid worst-case conditions.
- Be sure adequate redundancy exists.
- Identify and implement lines of defense:
  Preventive measures.
  Corrective measures.
  Sometimes, as a last resort, containment or, possibly, abandoning the process if the hazard is unacceptable.

It is important to note that the worst case should be something that is realistic, not something that is conceivable but which may be unreasonable, such as a chemical plant far from any airport being struck by a large airplane.

Reactive Chemicals Testing. Much reactive chemical information involves thermal stability and the determination of:

- The temperature at which an exothermic reaction starts.
- The rate of reaction as a function of temperature.
- Heat generated per unit of material.

The information can be obtained by a variety of types of laboratory tests and by thermodynamic calculations, although reliable kinetic data cannot usually be obtained solely by calculations.

Losses from Dust Explosions

Most organic solids, most metals, and some combustible inorganic salts can form explosive dust clouds. In order to have a dust explosion, it is necessary to satisfy certain conditions:

- Suitably sized dust particles.
- Sufficient source of ignition energy.
- Dust concentration within explosive limits.
- Explosible dust.

If an explosive dust in air that meets the above criteria occurs in a process, an explosion should be considered as inevitable. The process designer should take into account the possibility of dust explosions and design accordingly.

In dust explosions the combustion process is very rapid. The flame speed is high compared to that in gas deflagrations. Detonations normally do not occur in dust explosions in industrial plants.

The sequence of events in a serious industrial dust explosion is often as follows:

1. A primary explosion occurs in part of a plant, causing an air disturbance.
2. The air disturbance disperses dust and causes a secondary explosion which is often more destructive than the primary explosion.

If the occurrence of a flammable (explosive) dust is inevitable in a particular process, several design alternatives or combinations of alternatives are available:

- Containment (maximum pressure of a dust explosion is usually below 120 to 150 psig).
- Explosion venting to a safe place.
- Inerting (most organic dusts are nonflammable in atmospheres containing less than about 10% oxygen).
- Suppression.

A fundamental solution to the dust explosion problem is to use a wet process so that dust suspensions do not occur at all. However, the process must be wet enough to be effective. Some dusts with a high moisture content can still be ignited.

Dust concentrations in major equipment may be designed to be below the lower flammable limit, but this often cannot be depended on in actual operation. Dust concentrations cannot be safely designed to be above an upper flammable limit because such a limit is ill-defined.[2]

For a large number of flammable dusts, the lower explosion limit lies between 0.02 and 0.06 kg/m.[3] The upper explosion limit is in the range of 2 to 6 kg/m,[3] but this number is of limited importance.

A small amount of flammable gas or vapor mixed in with a flammable dust can cause an explosive mixture to be formed even if both are at concentrations below the explosive range by themselves. These mixtures are called "hybrid" mixtures. The ignition energy to ignite a hybrid mixture is often less than that required for the flammable dust by itself.

Venting is only suitable if there is a safe discharge for the material vented. Whenever an explosion relief venting device is activated, it may be expected that a tongue of flame containing some unburned dust will first be ejected. The unburned dust will be ignited as it flows out the vent and can produce a large fireball that will extend outward, upward, and downward from the vent. It is essential for protection of personnel that venting is to an

open place not used by people. If a duct must be used, the explosion pressure in the enclosure will be increased considerably. Therefore, particular attention must be paid to the design of the enclosure in which the explosion could take place.

The NFPA 68 guide issued in 1988[7] has nomographs which can be used to select relief areas required for combustible dusts when test data on the dusts are available. The nomographs in NFPA 68 are by far the preferred way to design dust exlosion relief devices.

Relief venting to reduce dust explosion pressure requires the equipment to be protected to have a certain minimum strength. If the enclosure strength is too low, the enclosure will be damaged or destroyed before the explosion relief device can function. NFPA 68[7] states that the strength of the enclosure should exceed the vent relief pressure by at least 0.35 psi. For industrial equipment such as dryers and baghouses, it is often desirable to have considerably more strength built into the structure to reduce the size of the vent area required. Also, the supporting structure for the enclosure must be strong enough to withstand any reaction forces developed as a result of operation of the vent.

Inerting is a very good preventive measure against dust explosions. The maximum oxygen concentration at which dust explosions are "just not possible" cannot be predicted accurately, as it depends on the nature of the combustible material; testing is usually required. It has been found that in an atmosphere of 10 percent oxygen and 90 percent nitrogen, most combustible organic dusts are no longer explosive. To allow a safety margin, it is good industrial practice to maintain oxygen concentrations below 8 percent. For metal dusts, the allowable oxygen content is about 4 percent.[7]

Inerting leads to the possibility of asphyxiation by operating personnel if they should be exposed to the inert gas. Strict precautions must be taken to prevent exposure of personnel to inerting atmospheres.

Explosion suppression systems are designed to prevent the creation of unacceptably high pressure by explosions within enclosures that are not designed to withstand the maximum explosion pressure.[8] They can protect process plants against damage and also protect operating personnel in the area. Explosion suppression systems restrict and confine the flames in a very early stage of the explosion. Suppression systems require more maintenance than do relief venting devices. Explosion suppression systems are made by only a few manufacturers and are quite expensive. This may be the reason why this type of safeguard has not been as widely used in industry as one might expect, although its effectiveness has been proved by much practical experience.

Explosion suppression is a proven technology and should be considered as a candidate for explosion protection. The NFPA has published a standard reference on explosion-suppression protection.[9] Manufacturers should be consulted on design, installation, and maintenance.

Even with explosion suppression, it is common for the explosion pressure to reach one atmosphere before it is suppressed. The added pressure surge from the injection of the suppressing agent must also be considered. Therefore, sufficient mechanical strength is always required for enclosures protected by explosion suppression.

Substitution of Less Hazardous Materials

It may be possible to substitute a less hazardous material for a hazardous product. For example, bleaching powder can be used in swimming pools instead of chlorine.[4] Benzoyl peroxide, an initiator used in polymerization reactions, is available as a paste in water, which makes it much less shock-sensitive than the dry form. Other substitutions that have been used to make transportation, storage, and processing safer include:

- Shipping ethylene dibromide instead of bromine.
- Shipping ethyl benzene instead of ethylene.
- Storing and shipping chlorinated hydrocarbons instead of chlorine.

- Storing and shipping methanol instead of liquefied methane.
- Replacing flammable refrigerants by halogenated hydrocarbons that are known to have acceptable environmental effects.
- Storing and shipping carbon tetrachloride instead of anhydrous hydrochloric acid. The CCl_4 is burned with supplemental fuel to make HCl on demand at the user's site.
- Using magnesium hydroxide slurry to control pH instead of concentrated sodium hydroxide solutions, which are corrosive to humans and relatively hazardous to handle.
- Using pellets of flammable solids instead of finely divided solids to reduce dust explosion problems.

The use of substitutes may appear to be more costly. The added safety provided by substitutes may make their use worthwhile and can in some cases actually lower the true cost of the project when the overall impact on the process, surrounding areas, and shipping is considered.

Substitutes should be employed only if it is known that overall risk will be reduced. Inadequately tested processes may introduce unrecognized health, safety, and environmental problems.

Catastrophic Failure of Engineering Materials[10–12]

Uniform corrosion of metals can usually be predicted from lab tests or experience. Corrosion allowances which will require thicker metal can be called for in the design of equipment when uniform corrosion rates are expected.

The most important materials failure to avoid in the design of metal equipment is *sudden catastrophic failure.* This occurs when the material fractures under impulse instead of bending. Catastrophic failure can cause complete destruction of piping or equipment, and can result in explosions, huge spills, and consequent fires. Causes of some of the more common types of catastrophic failures are:

- Low temperature brittleness.
- Stress corrosion cracking.
- Hydrogen embrittlement.
- High penetration rates involving pitting and corrosion.
- Fatigue failure.
- Creep.
- Mechanical shock.
- Thermal shock.
- High rates of temperature change in brittle materials.
- Zinc embrittlement of stainless steel.
- Caustic embrittlement.
- Nitrate stress corrosion.

Redundant Instrumentation and Control Systems[13]

Computer-controlled chemical plants are becoming the rule rather than the exception. As a result, it is possible to measure more variables and get more process information than ever, and chemical plants can be made "inherently safer" than ever before. However, it must be kept in mind that instruments and control components *will* fail. It is not a question of *if* they will fail, but *when* they will fail, and what the consequences will be. Therefore, the question of redundancy must be thoroughly considered. *The system must be designed so that when failure occurs, the plant is still safe.*

Redundant measurement means obtaining the same process information with two like measurements or two measurements using different principles. Redundant measurements can be *calculated* or *inferred* measurements. Two like measurements would be two pressure transmitters, two temperature measurements, two level measurements, and so on. An example of inferred measurement would be measuring temperature in a boiler and, using a pressure measurement and vapor pressure tables, determining the pressure to check an actual temperature measurement.

A continuous analog signal that is continuously monitored by a digital computer is generally preferable to a single point or single

switch such as a high level switch or high pressure switch. A continuous analog measurement can give valuable information about what the value is *now* and can be used to compute values or compare with other measurements. Analog inputs may be visual, and one can see what the set point is and what the actual value is. The software security system should determine who changes set points, and should not be easy to defeat.

A single point (digital) signal only determines whether switch contacts are open or not. It can indicate that something has happened, but not that it is going to happen. It cannot provide information to anticipate a problem that may be building up or a history about why the problem happened. Single point signals are easy to defeat. Some single point measurements are necessary, such as fire eyes, backup high level switches, and so on.

As a rule, it is best to avoid:

- Both pressure transmitters on the same tap.
- Both temperature measuring devices in the same well.
- Both level transmitters on the same tap or equalizing line.
- Any two measurements installed so that the same problem can cause a loss of both measurements.

It is a good idea to use devices that use different principles to measure the same variable, if possible.

An alarm should sound any time redundant inputs disagree. In many cases, the operating personnel will have to decide what to do. In some cases the computer control system will have to decide by itself what to do if redundant inputs disagree.

The more hazardous the process, the more it is necessary to use multiple sensors for flow, temperature, pressure, and other variables.

Since it must be assumed that all measuring devices will fail, they should fail to an alarm state. If a device fails to a non-alarm condition, there can be serious problems. If a device fails to an alarm condition, but there is really not an alarm condition, it is also serious, but generally not as serious as if it fails to a non-alarm condition, which can provide a false sense of security.

Usually it is assumed that two devices measuring the same thing will not fail independently at the same time. If this is assumed, one can consider the effects of different levels of redundancy:

Number of Inputs	*Consequence*
One	Failure provides no information on whether there is an alarm condition or not.
Two	Failure of one device shows that there is a disagreement, but without more information, it cannot be determined whether there is an alarm condition or not. More information is needed; the operator could "vote" if there is time.
Three	Failure of one device leaves two that work; there should be no ambiguity on whether there is an alarm condition or not.

Pressure Relief Systems

The design of relief systems involves, in general, the following steps:

1. Generate scenario. What could reasonably happen that could cause high pressures? This could be fire, runaway reactions, phase changes, generation of gases or vapors, leaks from high pressure sources, etc.
2. Calculate the duty requirements—the lb/hr of material that has to be vented, and its physical condition (temperature, pressure, ratio of vapor to liquid, physical properties). This is a rather involved calculational procedure.
3. Calculate the relief area required based on the duty, inlet and outlet piping, and downstream equipment. This is also a rather involved calculational procedure.
4. Choose the relief device to be specified from vendor information.

A group of chemical companies joined together in 1976 to investigate emergency relief systems. This later resulted in the formation of The Design Institute for Emergency Relief Systems (DIERS), a consortium

of 29 companies under the auspices of the AIChE. DIERS was funded with $1.6 million to test existing methods for emergency relief system design and to "fill in the gaps" in technology in this area, especially in the design of emergency relief systems to handle runaway reactions.[14] DIERS completed contract work and disbanded in 1984.

Huff was the first to publish details of a comprehensive two-phase flow computational method for sizing emergency relief devices, which, with refinements, has been in use for over a decade.[15-18] The most significant theoretical and experimental finding of the DIERS program was the ease with which two-phase vapor–liquid flow can occur during an emergency relief situation. *The occurrence of two-phase flow during runaway reaction relief almost always requires a larger relief system than does single-phase flow vapor venting.* The required area for two-phase flow venting can be from two to much more than two times larger to provide adequate relief than if vapor-only venting occurs.[15] Failure to recognize this can result in drastically undersized relief systems that will not provide the intended protection.

Two-phase vapor–liquid flow of the type that can affect relief system design occurs as a result of vaporization and gas generation during a runaway reaction or in many liquid systems subjected to fire (especially tanks that are nearly full). Boiling can take place throughout the entire volume of liquid, not just at the surface. Trapped bubbles, retarded by viscosity and the nature of the fluid, reduce the effective density of the fluid and cause the liquid surface to be raised. When it reaches the height of the relief device, two-phase flow results. Fauske and Leung[19] described test equipment that can be used to help determine the design of pressure relief systems for runaway reactions that often result in two- or three-phase flow.

Safe and Rapid Isolation of Piping Systems and Equipment

It should be possible to easily isolate fluids in equipment and piping when potentially dangerous situations occur. This can be done using emergency block valves, or EBVs. An EBV is a manually or remotely actuated protective device which should be used to provide manual or remote shutoff of uncontrolled gas or liquid flow releases. EBVs can be used to isolate a vessel or other equipment, or an entire unit operation. Manual valves are often used on piping at block limits where it is unlikely that there would be a hazard to personnel if an accident occurs. Remotely controlled EBVs are recommended on tanks and on piping in areas where it may be hazardous for personnel in case of an accident, or where quick response may be necessary.

EBVs used on tanks should be as close as possible to the tank flange and not in the piping away from the tank. In cases where EBVs may be exposed to fire, the valve and valve operating mechanism must be fire-safe.

In one design case involving large quantities of highly flammable materials, a HAZOP showed that suitably located EBVs were one of the most important features that could be installed to reduce the possibility of loss of containment and serious fire and explosions.

Piping, Gaskets, and Valves

Piping.[20] All-welded pipe and flanges should be used in "inherently safer" chemical plants. Since flanges are a potential source of leaks, as few flanges as possible should be used. One, of course, has to be realistic: if it is necessary to clean out pipes, flanges must be provided at appropriate places to make it possible. Also, enough flanges must be provided to make maintenance and installation of new equipment reasonably easy.

Threaded piping should be avoided for toxic and flammable materials. It is very difficult to make threaded fittings leakproof, especially with alloys such as stainless steel. Where threaded piping is necessary, use schedule 80 pipe as a minimum. Pipe nipples should never be less than schedule 80.

Pipe support design should be given special attention. It may be desirable to increase pipe diameter to provide more pipe strength and rigidity and make it possible to have greater

distance between supports. Normally, in chemical plants, it is not desirable to use piping less than one-half inch in diameter and preferably not less than one inch in diameter, even if the flow requirements permit a smaller pipe, except for special cases. Pipe smaller than one-half inch has insufficient strength and rigidity to be supported at reasonable intervals. Tubing, properly supported, should normally be used for anything smaller than one-half inch. Tubing is not as fragile as pipe in small sizes. It can be bent so that the number of fittings required is reduced. If it is necessary to use smaller pipe, or small tubing, special provisions should be made for its support and mechanical protection. Also, consideration should be given to using schedule 80 or schedule 160 pipe if small pipe is required to provide extra mechanical strength, even if the fluid pressure does not require it.

Gaskets.[21] Gaskets are among the weakest elements of most chemical plants. Blown-out or leaky gaskets have been implicated in many serious incidents. A leak at a flange can have a torch effect if it is ignited. A fire of this type was considered as a possible cause of the Flixborough disaster.[2] (See below under "Case Histories.")

Modern technology makes it possible to greatly reduce the incidence of gasket failure by the use of spiral wound gaskets. These are sold by several manufacturers, including Flexitallic, Parker Spirotallic, Garlock, and Lamons. A spiral wound gasket is considered the safest gasket type available because of the metal inner and outer rings which contain the filler material at the ID and OD. This sturdy construction of a spiral wound gasket does not permit blowout of the gasket material, which is the potential for gaskets made of sheet material.

Bolting with spiral wound gaskets is very important. *Use of plain carbon steel bolts (such as A307 Grade B) with spiral wound gaskets is not permitted.* They are not strong enough. High strength alloy bolts such as A193-B7 (contains Cr and Mo) with A194 heavy hex nuts should be used. To properly seal spiral wound gaskets, it is necessary to tighten the bolts to specified torque limits, which are generally higher than those of conventional gaskets. Compared to conventional gaskets, spiral wound gaskets require better flange finish, heavier flanges, and better flange alignment.

Valves. It is desirable and "inherently safer" to use firesafe valves whenever it is necessary to isolate flammable or combustible fluids in a pipeline or a tank or other type of equipment. Firesafe valves should be considered for handling most fluids that are highly flammable, highly toxic, or highly corrosive and that cannot be allowed to escape into the environment.

Firesafe valves should be used to isolate reactors, storage vessels, and pipelines. They can be used wherever EBVs are required. With the increase in popularity of automated plants, quarter-turn valves are very popular and are used in most installations. The only common quarter-turn valves that are available as completely firesafe valves are ball valves and high-performance butterfly valves. For special purposes, there are other special firesafe valves.

Plastic Pipe and Plastic-Lined Pipe. Plastic-lined pipe is excellent for many uses, such as highly corrosive applications, or where sticking is a problem, and where ease of cleaning is a factor. It is often the cheapest alternative. However, if there is a fire, there may be "instant holes" at each flange because the plastic will melt away, leaving a gap. Therefore, plastic-lined pipe should not ordinarily be used for flammable materials that must be contained in case of a fire. An exception to this is a firesafe plastic-lined pipe system such as that made by the Resistoflex Corp., which provides a metal ring between each flange that will make plastic-lined pipe firesafe. The pipe will probably have to be replaced after a fire, but the contents of the pipe will be contained during a fire.

In general, all types of solid plastic or glass-reinforced plastic pipe should be avoided, if possible, with flammable liquids. Compared

to metal, plastic piping:

- Will melt and burn easier.
- Is more fragile and can be easily mechanically damaged.
- Is harder to support adequately.

Plastic pipe should be used with appropriate judgment.

Avoidance of Inherently Unsafe Equipment

Some equipment items can be regarded as "inherently unsafe" for use in flammable or toxic service and should be avoided if possible. These items include a variety of devices.

Glass and transparent devices, such as sight glasses, bull's-eyes, sightports, rotameters, and glass and transparent plastic piping and fittings, may be hazardous. Glass devices are sensitive to heat and shock. Transparent plastic devices may be resistant to shock, but are not resistant to high temperatures. If they fail in hazardous service, severe property damage and personnel injury can result. The guidelines to follow are:

1. If broken, would they release flammable material?
2. If broken, would they expose personnel to toxic or corrosive materials?

Flexible or expansion joints in piping should be eliminated wherever possible. Flexible joints and expansion joints are any corrugated or flexible transition devices designed to minimize or isolate the effects of:

- Thermal expansion
- Vibration
- Differential settling
- Pumping surges
- Wear
- Load stresses
- Other unusual conditions

The need for flexible joints sometimes can be eliminated by proper design so that solid pipe will be able to handle misalignment, thermal expansion, and so on, by bending slightly. In many cases electronic load cells can be used to accurately weigh large reactors and process tanks that may have pipes attached to them, with no flexible or expansion joints. This is done by cantilevering the pipes to the equipment using sufficient runs of straight horizontal unsupported pipe with 90° elbows to take up movements and vibration without interfering significantly with the operation of the load cells.

Flexible joints should not be used as a correction for piping errors or to correct misalignment.

Pumps for Hazardous Service

A wide variety of excellent pumps are available in the chemical industry. It is sometimes a problem to choose the best from the large number available. This discussion will be limited to centrifugal pumps. Assuming that one has sized the pump, decided on a centrifugal pump, and chosen a suitable list of vendors, the main choices involve (1) metallurgy (2) seal-less pumps versus conventional centrifugal pumps, and (3) type of seals for conventional pumps.

Metallurgy.[22] Cast iron should not be used for flammable or hazardous service. The minimum metallurgy for centrifugal pumps for hazardous or flammable materials is cast ductile iron, type ASTM A 395, having an ultimate tensile strength of about 60,000 psi.

Seal-less Pumps.[23] The most common maintenance problem with centrifugal pumps is with the seals. Seals can be essentially eliminated with seal-less pumps. Seal-less pumps are manufactured in two basic types: canned motor and magnetic drive. Magnetic drive pumps have thicker "cans," which hold in the process fluid and the clearances between the internal rotor and "can" are greater compared to canned motor pumps. Seal-less pumps are becoming very popular and are widely used in the chemical industry. Mechanical seal problems account for most of the pump repairs in a chemical plant, with bearing

failures a distant second. The absence of an external motor (on canned pumps) and a seal is appealing to those experienced with mechanical seal pumps. However, it cannot be assumed that just because there is no seal, seal-less pumps are always safer than pumps with seals, even with the advanced technology now available in seal-less pumps. Seal-less pumps must be used with considerable caution when handling hazardous or flammable liquids. A mistreated seal-less pump can rupture with potentially serious results. The "can" can fail if valves on both sides of the pump are closed, and the fluid in the pump expands because of heating up from a cold condition with the pump off, or if the pump is started up under these conditions. If the pump is run dry for even a short time, the bearings can be ruined. Seal-less pumps, especially canned motor pumps, produce a significant amount of heat because nearly all the electrical energy lost in the system is absorbed by the fluid being pumped. *If this heat cannot be properly dissipated, the fluid will heat up with possibly severe consequences.* Considerable care must be used in installing a seal-less pump to be sure that misoperations cannot occur.

Properly installed and maintained, seal-less pumps, especially magnetic drive pumps, offer an economical and safe way to minimize hazards and leaks of hazardous liquids.

TECHNICAL MANAGEMENT OF CHEMICAL PROCESS SAFETY

Although understanding and using sound technology is important in designing and operating safe process plants, technology is not enough. As the chemical process industries have developed more sophisticated ways to improve process safety, safety management systems have been introduced to make process safety engineering activities more effective. The following is a brief summary of a recent approach to the technical management of chemical process safety.[24]

Twelve process safety elements have been identified as important in the context of plant design, construction, start-up, operation, maintenance, modification, and decommissioning. This does not include personnel safety, transportation issues, chronic releases to the environment, or community response, which are separate and important issues. Process safety management must deal with each element. Even the best companies, with the best safety records, have room for improvement.

1. *Accountability.* Accountability begins with a clear, explicit and reasonably specific statement of a company's expectations, objectives, and goals. Example: "Process safety audits must be conducted based on the relative risk involved."

2. *Process Knowledge and Documentation.* Capturing process knowledge is a foundation upon which many aspects of a process safety program are built. This is sometimes referred to as "company memory." Preserving this knowledge and making it available within a company are important for process safety for a number of reasons, including:

- Keeping a record of design conditions and materials of construction to help ensure that operations remain faithful to the original intent.
- Providing a basis for understanding how and why the process should be operated a certain way.
- Making it possible to pass information from older to younger workers.
- Providing a baseline for process changes.
- Recording causes of incidents and accidents and corrective action for future guidance.

3. *Process Safety Reviews for Capital Projects.* The need for process safety reviews of capital projects, whether new or revised facilities, has been a recognized feature of engineering organizations. These reviews must be comprehensive and systematic. It is best to do reviews early to avoid costly modifications later. Not only do process hazards need to be identified and addressed by the design, but broader issues should be considered, such

as:

- Is the company prepared to accept the hazards and risks of this project?
- Is the location appropriate?

4. *Process Risk Management.* Process risk management involves the systematic identification, evaluation, and control of potential losses that may arise in existing operating facilities from events such as fires, explosions, toxic releases, runaway reactions, or natural disasters. If risks are not identified, they cannot be considered. Whether resulting losses are measured in terms of direct costs, impacts on employees and/or the public, property or environmental damage, lost business, or various liabilities, the possibility of experiencing such losses is considered a risk. When risks have been identified, it is possible to plan for their reduction, and it can be determined whether the remaining, or residual, risk is acceptable. Risk can never be entirely eliminated. It is the purpose of this component of risk management to manage the risks that remain after implementation of risk controls.

5. *Management of Change.* Changes to process facilities are necessary for many reasons. Change includes all modifications to equipment, procedures, and organizations that may affect process safety. Normal operation of a process should be within defined safe limits; operation outside of these limits should require review and approval. Replacement personnel should be trained before moves of people are made. There should be no difference between things to consider for changes intended to be permanent and changes that have a limited life, that is, are "temporary" in nature. A hazard will proceed to an incident whether the change is permanent or temporary. In any operation, situations arise that were not foreseen. To assure that these deviations from normal practice do not create unacceptable risks, it is important to have a variance procedure incorporated into the management system.

6. *Process and Equipment Integrity.* Equipment used to handle or process hazardous materials should control the risk of releases and other accidents. Management systems should ensure that equipment is designed properly and constructed and installed according to the design. Components of the process and equipment integrity element include: reliability engineering, materials of construction, inspection procedures, installation procedures, preventive maintenance, and maintenance procedures. It is very important that a program of tests and inspections be available to detect impending failures and mitigate their potential before they can develop into more serious failures.

7. *Incident Investigation.* Incidents can be defined broadly as unplanned events with undesirable consequences. Incident investigation is the management process by which the underlying causes of incidents are uncovered and steps are taken to prevent similar incidents. Almost always, it has been found that process safety incidents are a result of some failure of the management system, which, had it functioned properly, could have prevented the incident. Incident investigations should identify all underlying causes and management system failures. It is not enough to discover the immediate cause, such as operator error, but the investigation should go deeper and find out why the operator made the error and possibly discover training and equipment deficiencies. The incident investigation should report facts and conclusions and make recommendations.

8. *Training and Performance.* Training is an essential part of any process safety management program. Proper training of personnel is an absolute requirement for keeping complex process equipment and machinery operating safely. Good training programs tell the student not just what to do, but also how to do it and why it is important. Teaching should be a high-priority task. There should be specific criteria for instructor selection. It is not enough to know the subject, although that is important; the teacher should have good presentation skills. The effectiveness of training should be evaluated regularly to determine the effectiveness of the instructors and the performance of the students.

9. *Human factors (Ergonomics).* Human factors or ergonomics refers to human characteristics that must be considered in designing technical systems and equipment so that they can be used safely and efficiently. Anything that increases the difficulty of an operator's job may result in errors if shortcuts are taken. Such shortcuts may work under normal conditions, but unacceptable in upset or abnormal conditions. The process safety management system should address the human–hardware interface, in both automated and manual processes. Safety reviews and HAZOP studies should include exmination of the human–hardware interface. Even in a nearly perfectly designed situation, operators will make occasional errors. Consideration should be given to involving specialized experts in human error assessment, especially in processes involving high potential risks and extensive operator control.

10. *Company Standards, Codes, and Regulations.* Company standards, codes, and regulations are intended to communicate minimum acceptable safe practice and to assure that all operating locations within the company share a common approach to process safety. All U.S. chemical plants are also subject to federal government regulations, and in some states there are specific environmental and toxic laws that apply. Plants in many countries outside the U.S. have specific laws that apply to those countries. Many companies also have internal standards and guidelines to ensure consistency in decision-making by design engineers and plant personnel. There should be clear documentation so that it is known which codes, standards, and guidelines are to be followed.

11. *Audits and Corrective Actions.* Audits are methodical, independent, and usually periodic examinations of local installations, procedures, and practices. Audits help assure compliance with a sound process safety program and that risks are being properly managed. The frequency of audits should be determined by the hazards of the facility. Corrective actions are the steps taken by a company in response to recognition of a process safety deficiency, either through an audit or by other means. It is important that an audit team have expertise in facility operations, safety disciplines such as fire protection, and management systems. Corrective action includes the process of addressing identified deficiencies, weaknesses, or vulnerabilities. It also includes processes for corrective action planning and follow-up.

12. *Enhancement of Process Safety Knowledge.* Organizations with strong process safety programs can contribute to advancing the latest process safety technology. They should share nonproprietary results of safety research and support process safety–oriented research and development programs of professional and trade associations and colleges. Improved process safety knowledge can produce a competitive advantage by improved yields, better quality, increased productivity, and less downtime. Consideration should be given to encouraging technical staff participation in professional and trade association programs. The Center for Chemical Process Safety of the American Institute of Chemical Engineers, which is supported by many large chemical and chemical-related companies, is an example of a technical group that sponsors conferences, training programs, and publications in the area of chemical process safety.

COMBUSTION HAZARDS

Introduction

The enchanting flame has held a special mystery and charm the world over for thousands of years. According to Greek myth, Prometheus the Titan stole fire from the heavens and gave it to mortals—an act for which he was swiftly punished. Early people made use of it anyway. Soon the ancients came to regard fire as one of the basic elements of the world. It has since become the familiar sign of the hearth and a mark of youth and blood, as well as the object of intense curiosity and scientific investigation.

Suitably restrained, fire is of great benefit; unchecked or uncontrolled, it can cause immense damage. We respond to it with a

powerful fascination coupled with an inbred respect and fear. A good servant but a bad master is Thoreau's "most tolerable third party."[25]

Fire[26]

Fire or combustion is normally the result of fuel and oxygen coming together in suitable proportions and with a source of heat. The consumption of a material by a fire is a chemical reaction in which the heated substance combines with oxygen. Heat, light, smoke, and products of combustion are generated. The net production of heat by a fire involves both heat-producing and heat-absorbing reactions, with more heat being produced than is absorbed. Energy in the form of heat is required:

1. To produce vapors and gases by vaporization or decomposition of solids and liquids. Actual combustion usually involves gases or vapors intimately mixed with oxygen molecules.
2. To energize the molecules of oxygen and flammable vapors into combining with one another and so initiating a chemical reaction.

The amount of energy required to cause combustion varies greatly. Hydrogen and carbon disulfide can be ignited by tiny sparks, or simply may be ignited by static generated as the gases or vapors discharge from pipes into air. Other materials, such as methylene chloride, require such large amounts of energy to be ignited that they sometimes are considered nonflammable. Fire also can result from the combining of such oxidizers as chlorine and various hydrocarbon vapors; oxygen is not required for a fire to take place.

There are exceptions to the general rule that a solid must vaporize or decompose to combine with oxygen; some finely divided materials such as aluminum powder and iron power can burn, and it is generally accepted that they do not vaporize appreciably before burning.

Products of Combustion. Heat, light, smoke, and asphyxiating toxic gases are produced by fire. In a hot, well-ventilated fire, combustion usually is nearly complete. Nearly all the carbon is converted to carbon dioxide, and all the hydrogen to steam, and oxides of various other elements such as sulfur and nitrogen are produced.

This is not the case in most fires, where some of the intermediate products, formed when large complex molecules are broken up, persist. Examples are hydrogen cyanide from wool and silk; acrolein from vegetable oils; acetic acid from timber or paper; and carbon or carbon monoxide from the incomplete combustion of carbonaceous materials. As the fire develops and becomes hotter, many of these intermediates, which are often toxic, are destroyed (e.g., hydrogen cyanide is decomposed at a significant rate at 538°C).

Small airborne particles of partially burnt carbonaceous materials form smoke, which is often thickened by steam, when there is only partial combustion of fuel.

Solids. Ordinarily, combustible solids do not combine directly with oxygen when they burn. They give off vapor and gaseous decomposition products when they are heated, and it is the vapors or gases that actually burn in the characteristic form of flames. Thus, before a solid can be ignited, it usually must be heated sufficiently for it to give off flammable concentrations of vapors. Glowing, which is combustion in the solid state, is characteristic of materials in the final stages of a fire's decay when flammable gases have been burned away, or when the production of gases and vapors has been suppressed.

Solids with larger surface areas, in relation to their volume, burn more readily than those that are more compact when exposed to heat and oxygen in the air. Common materials such as textiles in the form of fibers or fabrics, foamed rubber, foamed plastics, thin sheets of plastic, paper, corrugated cardboard, combustible dusts, dry grass and twigs, and wood shavings are examples of materials with large surface areas in relation to their volume. In a well-established fire, materials with relatively

small surface areas, such as chunks of coal or logs, burn readily.

Combustion is self-propagating; burning materials produce heat which causes more of the solid to produce flammable vapors until either the fuel or oxygen is exhausted, or until the fire is extinguished in some other way.

Dusts. Most combustible solids can produce combustible dusts. Combustible dusts are particularly hazardous; they have a very high surface area to volume ratio. When finely divided as powders or dusts, solids burn quite differently from the original material in the bulk. Dust and fiber deposits can spread fire across a room or along a ledge or roof beam very quickly. Accumulations of dust can smoulder slowly for long periods, giving little indication that combustion has started until the fire suddenly flares up, possibly when no one suspects a problem.

Many combustible dusts produced by industrial processes are explosible when they are suspended as a cloud in air. Even a spark may be sufficient to ignite them. After ignition, flame spreads rapidly through the dust cloud as successive layers are heated to ignition temperature. The hot gases expand and produce pressure waves that travel ahead of the flame. Any dust lying on surfaces in the path of the pressure waves will be thrown into the air and could cause a secondary explosion more violent and extensive than the first.

Liquids. A vapor has to be produced at the surface of a liquid before it will burn. Many common liquids give off a flammable concentration of vapor in air without being heated, sometimes at well below room temperature. Gasoline, for example, gives off ignitable vapors above about $-40°C$, depending on the blend. The vapors are easily ignited by a small spark or flame. Other liquids, such as fuel oil and kerosene, need to be heated until sufficient vapor is produced.

Many liquids can be formed into mists that will burn at temperatures where the vapor pressure is insufficient to produce a flammable mixture of the vapor and air.

For any flammable vapor there are maximum and minimum concentrations of the vapor in air beyond which it cannot burn. When the mixture of vapor in air is too weak, there is insufficient fuel for burning; when the mixture is too strong, there is insufficient oxygen for burning.

If the density of a flammable vapor is greater than that of air, as is normally the case, flammable concentrations may collect at low levels, such as at floor level or in basements, and can travel a considerable distance to a source of ignition, from which flames will then flash back.

Gases. Flammable gases usually are very easily ignited if mixed with air. Flammable gases often are stored under pressure, in some cases as a liquid. Even small leaks of a liquefied flammable gas form relatively large quantities of gas, which is ready for combustion.

The Fire Triangle. The well-known "fire triangle" (see Fig. 5.1) is used to represent the three conditions necessary for a fire:

1. Fuel.
2. Oxidizer: oxygen or other gaseous oxidizer such as chlorine; or liquid oxidizer such as bromine; or solid oxidizer such as sodium bromate.
3. Energy, usually in the form of heat.

If one of the conditions in the fire triangle is missing, fire does not occur; if one is removed, fire is extinguished. Usually a fire

Fig. 5.1. The fire triangle.

occurs when a source of heat contacts a combustible material in air, and then the heat is supplied by the combustion process itself.

The fire triangle indicates how fires may be fought or prevented:

1. Cut off or remove the fuel.
2. Remove the heat or energy—usually by putting water on the fire.
3. Remove the supply of oxygen—usually by foam or inert gas.

Flammability

No single factor defines flammability, but some relevant parameters include:

1. Flash point—often considered the main index of flammability; low flash points usually mean increased flammability.
2. Flammability limits; wide limits mean increased flammability.
3. Autoignition temperature; low temperature means increased flammability.
4. Ignition energy; low ignition energy means increased flammability.
5. Burning velocity; high velocity means increased flammability.

A combustion process is an exothermic reaction initiated by a source of ignition that produces more energy than it consumes. The speed at which the reaction proceeds through the mixture of reactants depends on the concentration of the flammable gas or vapor. This speed is lower at higher ("rich") as well as at lower ("lean") concentrations of the flammable gas than at the stoichiometric mixture. There are lower and upper limits beyond which the reaction cannot propagate through the gas mixture on its own. Some definitions follow:

Flammability limits. The range of flammable vapor–air or gas–air mixtures between the upper and lower flammable limits. Flammability limits are usually expressed in volume percent. Flammability limits are affected by pressure, temperature, direction of flame propagation, oxygen content, type of inerts, and other factors. The precise values depend on the test method.

Upper flammability limit. The maximum concentration of vapor or gas in air above which propagation of flame does not occur on contact with a source of ignition.

Lower flammability limit. The minimum concentration of vapor or gas in air or oxygen below which propagation of flame does not occur with a source of ignition.

The concentrations at the lower and upper flammability limits are roughly 50 percent and 200 to 400 percent, respectively, of the stoichiometric mixture. The maximum flammability usually (not always) occurs at the stoichiometric mixture for combustion.[2,27]

Flammable limits for mixtures of flammable gases and vapors. For mixtures of several flammable gases and vapors, the flammable limits can be estimated by application of Le Chatelier's equation, if the flammable limits of the components are known.[2]

$$L = \frac{1}{\sum_{i=1}^{n} \frac{y_i}{L_i}} \qquad U = \frac{1}{\sum_{i=1}^{n} \frac{y_i}{U_i}}$$

where:

L = lower flammability limit of the fuel mixture, vol. %.
L_i = lower flammability limit of fuel component i, vol. %.
U = upper flammability limit of the fuel mixture, vol. %.
U_i = upper flammability limit of fuel component i, vol. %.
y_i = concentration of fuel component i, vol. %.

This equation is empirical and is not universally applicable, but is useful and a reasonable approximation when actual mixture data are not available.

It is possible for a mixture to be flammable even though the concentration of each constituent is less than its lower limit.

Methods of measurement. Flammability limits are determined by measuring the volume percent of a flammable gas in an oxidizing gas that will form a flammable mixture, thus identifying the lower and upper flammable limits as well as the critical oxygen concentration (the minimum oxidizer concentration that can be used to support combustion).

Uniformity of lower limits on a mass basis. Concentrations of vapors and gases usually are reported in volume percent. As molecular weight increases, the lower limit usually decreases. On a mass basis, the lower limits for hydrocarbons are fairly uniform at about 45 mg/L at 0°C and one atm. Many alcohols and oxygen-containing compounds have higher values; for example, on a mass basis, ethyl alcohol in air has a lower limit of 70 mg/L.[27]

Effect of temperature on flammable limits. The higher the temperature is at the moment of ignition, the more easily the combustion reaction will propagate. Therefore, the reference temperature (initial temperature) of the flammable mixture must be stated when flammable limits are quoted. There are not a lot of data for flammable limits under different conditions of initial temperature. The behavior of a particular mixture under different conditions of initial temperature usually must be determined by tests.

Burning in atmospheres enriched with oxygen. The flammability of a substance depends strongly on the partial pressure of oxygen in the atmosphere. Increasing oxygen content affects the lower flammability limit only slightly, but it has a large effect on the upper flammability limit. Increasing oxygen content affects the lower flammability limit only slightly, but it has a large effect on the upper flammability limit. Increasing oxygen content has a marked effect on the ignition temperature (reduces it) and the burning velocity (increases it).

At the lower explosive limits of gas–air mixtures, there is already an excess of oxygen for the combustion process. Replacing nitrogen by additional oxygen will influence this limit very little.[8]

Burning in chlorine. Chemically, oxygen is not the only oxidizing agent, though it is the most widely recognized and has been studied the most. Halogens are examples of oxidants that can react exothermically with conventional fuels and show combustion behavior. The applicability of flammability limits applies to substances that burn in chlorine. Chlorination reactions have many similarities to oxidation reactions. They tend not to be limited to thermodynamic equilibrium and often go to complete chlorination. The reactions are often highly exothermic. Chlorine, like oxygen, forms flammable mixtures with organic compounds.

Example. A chlorine–iron fire occurred in a chlorine pipeline, causing a chlorine gas release. Chlorine had liquefied in the lines because of the very cold weather, and the low spot was steam-traced. Steam had been taken from the wrong steam line, using 400 psig steam instead of 30 psig steam. The 400 psig steam was hot enough to initiate the reaction. This serves as a reminder that steel and chlorine can react. The allowable temperature for safe use depends upon the state of subdivision of the iron.

Burning in other oxidizable atmospheres. Flames can propagate in mixtures of oxide of nitrogen and other oxidizable substances. For example, Bodurtha[27] reports that the flammability limits for butane in nitric oxide are 7.5 percent (lower) and 12.5 percent (upper).

Flame quenching. Flame propagation is suppressed if the flammable mixture is held in a narrow space. There is a minimum diameter for apparatus used for determination of flammability limits. Below this diameter the flammable range measurements are narrower and inaccurate.

If the space is sufficiently narrow, flame propagation is suppressed completely. The largest diameter at which flame propagation

is suppressed is known as the quenching diameter. For an aperture of slotlike cross section there is critical slot width.

The term "quenching distance" sometimes is used as a general term covering both quenching diameter and critical slot width, and sometimes it means only the latter.

There is a maximum safe gap measured experimentally that will prevent the transmission of an explosion occurring within a container to a flammable mixture outside the container. These data refer to a stationary flame. If the gas flow is in the direction of the flame propagation, a smaller gap is needed to quench the flame. If the gas flow is in the opposite direction, a larger gap will provide quenching. If the gas velocity is high enough, the flame can stabilize at the constriction and cause local overheating. These quenching effects are important in the design of flame arrestors.

Heterogenous mixtures.[28] In industry, heterogeneous (poorly mixed) gas phase mixtures can lead to fires that normally would be totally unexpected. It is important to recognize that heterogeneous mixtures can ignite at concentrations that normally would be nonflammable if the mixture were homogeneous. For example, one liter of methane can form a flammable mixture with air at the top of a 100-liter container although the mixture only would contain 1.0 percent methane by volume if complete mixing occurred at room temperature, and the mixture would not be flammable. This is an important concept because "layering" can occur with any combustible gas or vapor in both stationary and flowing mixtures.

Heterogenous gas phase mixtures can lead to unexpected fires if a relatively small amount of flammable gas is placed in contact with a large amount of air without adequate mixing, even though the average concentration of flammable gas in the mixture is below the flammable limit. *Heterogeneous mixtures are always formed at least for a short time when two gases or vapors first are brought together.*

Effect of pressure. Flammability is affected by initial pressure. Normal variations of atmospheric pressure do not have any appreciable effect on flammability limits.

A decrease in pressure below atmospheric usually narrows the flammable range. When the pressure is reduced low enough, a flame or an explosion no longer can be propagated throughout the mixture.

An increase in pressure above atmospheric usually (not always) widens the flammability range, especially the upper limit.

Explosions in the absence of air. Gases with positive heats of formation can be decomposed explosively in the absence of air. Ethylene reacts explosively at elevated pressure, and acetylene reacts explosively at atmospheric pressure in large-diameter piping. Heats of formation for these materials are +52.3 and +227 kilojoules/g mol, respectively. Explosion prevention can be practiced by mixing decomposable gases with more stable diluents. For example, acetylene can be made nonexplosive at a pressure of 100 atmospheres by including 14.5 percent water vapor and 8 percent butane.

Ethylene oxide vapor will decompose explosively in the absence of oxygen or air under certain conditions when exposed to common sources of ignition if heated to high enough temperatures. One way to prevent the decomposition reaction is to use methane gas to blanket the ethylene oxide liquid. It has also been found that liquid ethylene oxide will undergo a deflagration in the absence of oxygen with a very rapid pressure increase if ignited at a temperature and pressure above a certain level. Fortunately, the conditions required for propagation of the decomposition of liquid phase ethylene oxide are outside the current normal handling and processing ranges for the pure liquid. Propagation has not been observed below 80°C at from 14 to 100 atm pressure.[29] Ethylene oxide also can undergo explosive condensation when catalyzed by a small amount of caustic.[30]

Inert Gases

The addition of inert gases to a mixture of flammable gases and air affects flammability limits. Carbon dioxide causes a greater narrowing of the flammable range than does nitrogen. Water vapor is an acceptable inert gas if the temperature is high enough to exclude much of the oxygen, which requires a temperature of 90 to 95°C. Because water vapor and carbon dioxide have a higher heat capacity than nitrogen, they are somewhat more effective as inerting agents than nitrogen. Some halogen-containing compounds also can be used for inerting materials at relatively low concentrations. An example of this is the use of Freon-12 (CCl_2F_2). Caution must be used with halohydrocarbons because of the possibility of the halocarbons themselves burning, especially at high pressures. Environmental considerations are making the use of halogenated hydrocarbons for inerting increasingly undesirable. Materials are being developed that are considered environmentally acceptable. Figures 5.2 and 5.3[28] show flammability envelopes for methane and *n*-hexane for various air–inert mixtures at 25°C and one atmosphere. All flammable envelopes are similar to Figs. 5.2 and 5.3 except in minor detail. The lower limit is virtually insensitive to added inerts. The upper limit, however, decreases linearly with added inert until the critical concentration of inert is reached beyond which no compositions are flammable. In these graphs, C_{st} means the stoichiometric composition.

Fig. 5.3. Limits of flammability of various *n*-hexane–inert gas–air mixtures at 25°C and atmospheric pressure. (*Courtesy Bureau of Mines.*)

Fig. 5.2. Limits of flammability of various methane–inert gas–air mixtures at 25°C and atmospheric pressure. (*Courtesy Bureau of Mines.*)

Figure 5.4 shows the minimum amount of nitrogen or carbon dioxide that must be added to air to suppress flammability of some common hydrocarbons.

Figure 5.5 shows the minimum oxygen content for combustion for some common flammable materials. The limits of flammability are dictated by the ability of a system to propagate a flame front. Propagation does not occur until the flame front reaches about

Fig. 5.4. Minimum inert gas content for suppression of flammability of selected substances in air. (*Data from Zabetakis 1965.*[28])

Fig. 5.5. Minimum oxygen content for combustion of selected substances. (*Data from Frank Bodurtha 1980.*[27])

1200 to 1400 K. As the typical terminal temperature for hydrocarbons at stoichiometric conditions is about 2300 K, it can be seen that having only one-half the fuel or oxidizer present will produce about one-half the flame temperature, which is too low to propagate flame.

A useful rule to remember is that the lower flammable limits of most flammable vapors are close to one-half the stoichiometric composition, which can be calculated. Another easy rule to remember is that about 10 percent oxygen or less in air (assuming the rest is mostly nitrogen) will not support combustion of most flammable hydrocarbon vapors.

The flammability limits of hydrocarbon type fuels in oxygen and inert gas atmospheres are a function of the inert gas and any fuel or oxygen in excess of that required by the stoichiometry of the combustion process. In systems where fuel content is fixed, inert material having a high heat capacity will be more effective at flame suppression than inert material having low heat capacity.

Many of the flammable limits reported in the literature are somewhat too narrow, and certain gas compositions regarded as being nonflammable are in fact flammable when given the proper set of circumstances. *In other words, take data on flammability limits from the literature with a grain of salt.* It is best not to design closely on the basis of most available data on flammability limits.

The use of inert gases can cause some serious hazards that must be recognized if inerts are to be used effectively and safely. Considerations in the use of inert gases include:

1. *An inert atmosphere can kill if a person breathes it.* Precautions should be taken to assure that personnel cannot be exposed to the breathing of inert atmospheres.

2. *Some products need at least a small amount of oxygen to be stored safely.* This includes styrene and some other vinyl monomers, which must have some oxygen in them to make the usual polymerization inhibitor for styrene (*t*-butyl catechol, or TBC), effective. If pure nitrogen, for example, is used to blanket styrene, the inhibitor will become ineffective. TBC customarily is added to styrene monomer to prevent polymer formation and oxidation degradation during shipment and subsequent storage; it functions as an antioxidant and prevents polymerization by reacting with oxidation products (free radicals in the monomer). If sufficient oxygen is present, polymerization is effectively prevented (at ambient temperatures); but in the absence of oxygen, polymerization will proceed at essentially the same rate as if no inhibitor were present. The styrene may polymerize and can undergo an uncontrolled exothermic reaction, which may generate high temperatures and pressures that can be very hazardous. The inhibitor level of styrene must be maintained above a minimum concentration at all times. The minimum concentration of TBC in styrene for storage is about 4 to 5 ppm.

3. *To be effective, inert atmospheres must be maintained within certain composition limits.* This requires the proper instrumentation and regular attention to the system.

4. *Inerting systems can be quite expensive and difficult to operate successfully.* Before the use of inert systems, alternatives should be explored, such as the use of nonflammable materials or operating well outside, preferably below, the flammability range.

Mists and Foams

If the temperature of a liquid is below its flash point, flammable concentrations of vapor cannot exist, but conditions still can exist for flammability if mists or foams are formed. A suspension of finely divided drops of a flammable liquid in air has many of the characteristics of a flammable gas–air mixture and can burn or explode. A mist may be produced by condensation of a saturated vapor or by mechanical atomization. Normally, the diameter of drops in a condensed mist is less than 0.01 mm, whereas in a mechanical spray it usually is greater than 0.1 mm.

The commonly accepted fallacy that liquids at temperatures below their flash points cannot give rise to flammable mixtures in air has led to numerous accidents. Flash

points are measured under stagnant conditions in carefully controlled laboratory experiments, but in the real world one works with a wide variety of dynamic conditions that can produce mists and foams.

Flammable mist–vapor–air mixtures may occur as the foam on a flammable liquid collapses.[28] Thus, when ignited, many foams can propagate flame. An additional hazard can arise from the production of foams by oxygen-enriched air at reduced pressures. Air confined over a liquid can become oxygen-enriched as pressure is reduced because oxygen is more soluble than nitrogen in most liquids. Thus, the presence of foams on combustible liquids is a potential explosion hazard.

The lower flammability limit for fine mists (<0.01 mm diameter) of hydrocarbons below their flash point, plus accompanying vapor, is about 48 grams of mist/cubic meter of air at 0°C and one atmosphere. Mist can occur in agitated vessels under some conditions, especially when an agitator blade is at or near the liquid–vapor interface in the vessel.

Work on condensed oil mists (drop diameter mostly less than 0.01 mm) has demonstrated that they have flammability characteristics similar to those the mixture would have if it were wholly in the vapor phase at the higher temperature necessary for vaporization. The flammability characteristics are affected by drop size. For larger drop sizes (above 0.01 mm) the lower limit of flammability decreases as drop diameter increases. For mists, the amount of inert gas needed to suppress flammability is about the same as that needed to suppress an equivalent vapor–air mixture of the same material if it were vaporized at a somewhat higher temperature.

A useful rule is that mists of flammable or combustible liquids in air can burn or explode at temperatures below their flash points.

Ignition

Flammable gases and vapors can be ignited by many sources. In the design and operation of processes, it is best not to base fire and explosion safety on the presumption that ignition sources have been excluded. Bodurtha[27] reported that of 318 natural gas fires and explosions, the sources of ignition of 28 percent were unknown. All reasonable measures should be taken to eliminate possible sources of ignition in areas in which flammable materials are handled.

Autoignition. If the temperature of a flammable gas–air mixture is raised in a uniformly heated apparatus, it eventually reaches a value at which combustion occurs in the bulk gas. This temperature is defined as the spontaneous ignition temperature (SIT) or autoignition temperature (AIT). The gas–air mixture that has the lowest ignition temperature is called by various names, such as minimum autoignition temperature, minimum spontaneous ignition temperature, and self-ignition temperature.[27] Usually the AIT reported in the literature is the minimum autoignition temperature.

The autoignition temperature of a substance depends on many factors, such as:

- Ignition delay
- Energy of ignition source
- Pressure
- Flow effects
- Surfaces
- Concentration of vapors
- Volume of container
- Oxygen content
- Catalytic materials
- Flow conditions

Thus, a specific autoignition temperature applies only to the experimental conditions employed in its determination. Usually the values quoted are obtained in clean laboratory equipment.

The autoignition temperature of a substance may be reduced below ideal laboratory conditions by as much as 100 to 200°C for surfaces that are insulated with certain types of insulation, or are contaminated by dust.

Mixtures that are fuel-rich or fuel-lean ignite at higher temperatures than do those of intermediate compositions. Also, in a homologous series of organic compounds,

Fig. 5.6. Autoignition temperatures of paraffin hydrocarbons at one atmosphere. (*Data from Bodurtha 1980.*[27])

the autoignition temperature decreases with increasing molecular weight, as shown in Fig. 5.6. (Data from Bodurtha[27]).

Ignition Delay. Ignition of a flammable mixture raised to or above the temperature at which spontaneous combustion occurs is not instantaneous; the time delay between the moment of exposure to high temperature and visible combustion is called the ignition delay. This time delay decreases as the ignition temperature increases. The time delay may be as little as a fraction of a second at higher temperatures, or several minutes close to the autoignition temperature.

Environmental Effects. It has been found that the AIT becomes lower with increasing vessel size in the range of 35 ml to 12 liters. An increase in pressure usually decreases AITs, and a decrease in pressure raises AITs. Usually oxygen enrichment of the air tends to decrease the minimum AIT, and a decrease in oxygen content increases the minimum AIT. Low-temperature oxidation can result in "cool flames," which may grow into ignition.

Catalytic Materials. Ignition may occur where the temperature is less than the minimum AIT. Catalytic materials, such as metal oxides, can promote oxidation on their surfaces, leading to a high local temperature and subsequent ignition. There is a recorded reactive chemical case[31] in which a solvent at 80°C was being pressurized with a gas phase consisting of a high oxygen concentration. The solvent has a flash point in oxygen of greater than 130°C and normally is considered not to be a flammability hazard. There was an ignition, causing the vessel to rupture its main gasket with major damage to the facility. It was found that a mist had been formed in the vessel by the agitator, and that the source of ignition probably was a trace of palladium catalyst remaining from a previous run.

From this incident, several important lessons can be learned:

1. Ignition of a flammable mixture can result from totally unexpected contamination by trace amounts of catalysts if the oxidizer and fuel are present.
2. Mists of oxidizable liquids may form that can burn or explode at temperatures outside the "normal" flammable range.

3. It can be dangerous to perform experiments with pure oxygen, or air enriched with oxygen, especially under pressure and at elevated temperatures, when oxidizable materials are present.
4. The real criterion regarding flammable mixtures in air should be whether a flammable atmosphere can exist under the given process conditions, rather than whether a flammable liquid is at a temperature below its flash point.

Cleaning Up Spills of Flammable or Combustible Liquids. It is customary to clean up small spills of many liquid materials with sand or other noncombustible absorbent material. Some absorbing agents, such as untreated clays and micas, will cause an exothermic reaction with some liquids, especially monomers, which might ignite the liquid if it is flammable or combustible. Before any material is provided to be used to soak up spills of oxidizable material, tests should be made to determine if the material can cause fires with potential spills.

Ignition Caused by Insulation. Ignition of combustible materials that have been absorbed into commonly used insulating materials is possible at temperatures lower than the AIT for nonabsorbed material. All oxidizable materials oxidize to some extent in air at ambient temperatures, usually at a very low rate. When an absorbent material is absorbed into insulation, it is "spread" over a large area, increasing its access to oxygen. Because the absorbent is an insulator, heat from oxidation is retained rather than dissipated, and the temperature will rise if the heat is produced faster than it can be dissipated. The rate of oxidation increases as the material temperature increases, which produces more heat, compounding the hazard. If the temperature rises enough, the material will ignite ("spontaneous combustion"). This is similar to the classic oily rag and wet haystack phenomenon, which has caused many fires in homes and on farms. In the wet haystack phenomenon, fermentation by microorganisms will create heat. Some air is necessary; too much air will remove too much heat to allow the combustion temperature to be reached. For equipment operating above about 200°C containing combustible liquids with high boiling points, insulation should be impervious to the material handled. To date, only a closed cell foamed glass provides the required degree of protection where oxidizable liquid materials are used above 200°C. Insulation based on glass fiber, silicate, or alumina materials is known to cause hazardous situations and should not be used in this service.

Laboratory tests and actual fires show that Dowtherm A® (a heat-transfer fluid consisting of a eutectic mixture of biphenyl oxide and biphenol) can be ignited if it is soaked in glass fiber insulation and in contact with air at temperatures considerably below the normal AIT. This is also true for stearic acid soaked in glass fiber insulation. Table 5.1 shows the reduction in AIT of Dowtherm A® and of steric acid soaked in glass fiber insulation.

Ignition of this type generally occurs only with materials having a high boiling point. Usually materials with low boiling points will vaporize and cannot remain soaked in hot insulation. There are exceptions. For example, ethylene oxide has a fairly low boiling point, but if it leaks into insulation, a polymer can be formed that has a high boiling point and can autoignite insulation at low temperatures.

Ignition Caused by Impact. Solids and liquids can be ignited by impact. Impact tests are made by having a weight fall freely through a known distance and impacting the sample. Impact can occur, for example, if containers are accidentally dropped. The

TABLE 5.1 Reduction in Autoignition Temperatures Caused by Liquids Soaking in Glass Fiber Insulation

Material	*Normal Autoignition Temperature, °C*	*Ignition Temp. in Glass Fiber Insulation, °C*
Dowtherm A®	621	260 to 290
Stearic acid	395	260 to 290

interpretation of the data from impact tests can be difficult.

Ignition Caused by Compression of Liquids. Liquids can be ignited by sudden compression. This can happen when there is water hammer caused by the pressure surge from quick-acting valves and from the compression in liquid pumps. Sudden compression can occur with liquids, for example, if a tank car is bumped rapidly and the liquid goes to one end very quickly, possibly trapping some vapor bubbles that compress and create local hot spots that can cause ignition.

Ignition Caused by Rubbing Friction. Solids can be ignited by frictional sources when rubbed against each other or against another material. The frictional heat produced may be enough to ignite other materials, such as lubricants, that are nearby. A common example of this occurs when bearings run hot, causing oil or grease to vaporize and possibly ignite.

Ignition Caused by Glancing Blows. Friction can cause ignition in other ways. Sparks may occur when two hard materials come in contact with each other in a glancing blow (the blows must be glancing to produce friction sparks). These kinds of sparks are not directly related to frictional impact. Hand and mechanical tools are the most likely sources of friction sparks that occur outside of equipment. The need for nonsparking tools is somewhat controversial; Bodurtha[27] states that it is extremely unlikely that anyone would be using tools in a flammable atmosphere, and it is usually more prudent to control the atmosphere than the tools. Sparkproof tools are not really sparkproof in all situations.

Ignition Caused by Static Electricity. Static electricity is a potential source of ignition wherever there is a flammable mixture of dusts or gases (see next section).

Ignition Caused by Compression of Gases. If a gas is compressed rapidly, its temperature will increase. Autoignition may occur if the temperature of the gas becomes high enough (this is more or less the principle of diesel engines).

An advancing piston of high-pressure gas can compress and heat trapped gas ahead of it. For a perfect gas, the temperature rise due to adiabatic compression is given by:

$$\frac{T_2}{T_1} = \left(\frac{P_2}{P_1}\right)^{(k-1)/k}$$

where T_1 and T_2 are the initial and final gas absolute temperatures, P_1 and P_2 are the initial and final absolute pressures, and k is the ratio of heat capacity at constant pressure to the heat capacity at constant volume. For air and many other diatomic gases, $k = 1.4$. Many hydrocarbons have k values of between 1.1 and 1.2. The value of k is a function of temperature and pressure.

Energy Levels for Ignition. If a flammable gas mixture is to be ignited by a local source of ignition, there is a minimum volume of mixture required to cause a continuing flame throughout the mixture.

For example, to ignite a methane–air mixture in a cold container, a hot patch of 18 mm^2 at 1000 to 1100°C is required in order to heat enough volume of gas to produce a continuing flame,[2] even though the auto-oxidation temperature for methane is 540°C. Ignition of a flammable gas–air mixture by electrical discharge can occur only if the electrical discharge is of sufficient energy.

Minimum Ignition Energy. There is a minimum ignition energy, which usually occurs near the stoichiometric mixture. The minimum ignition energy for some representative substances in air is shown in Fig. 5.7.[2] The energy required to cause ignition frequently is reported in millijoules (mJ). One joule is 0.24 calorie, so 1 mJ is 0.00024 calorie, which is a very small amount of energy.

A person typically has a capacitance of 200 picofarads, and if charged to 15 kilovolts could initiate a discharge of 22.5 mJ. This is enough to ignite many flammable mixtures. The energy in ordinary spark plugs is 20 to 30 mJ.

Fig. 5.7. Minimum ignition energy for selected substances. (*Lees 1980.*[2])

The hazard of an explosion should be minimized by avoiding flammable gas–air or dust–air mixtures in a plant. It is bad practice to rely solely on elimination of sources of ignition, as it is nearly impossible to ensure this.

Effect of Oxygen-Enriched Atmospheres. The minimum spark energy to cause ignition varies greatly with the amount of oxygen in oxygen-enriched air. Stull[30] showed that with a composition of 10 percent methane in air, about 0.5 mJ of spark energy is required to initiate a reaction at the lower flammable limit. If the air is enriched with oxygen, the minimum spark energy decreases. If the flammable material is combined with 100 percent oxygen, the spark energy required is only about one percent of the required energy in air at 21 percent oxygen! This demonstrates the extremely small amount of energy required to initiate the reaction, as well as the additional ease with which oxygen-enriched atmospheres are initiated. Table 5.2 compares initiation energies of some common substances in air and in pure oxygen.

Effect of pressure. An increase of pressure decreases the amount of energy required to cause ignition. In a mixture of propane, oxygen, and nitrogen, doubling the pressure decreases the minimum energy required to cause ignition by a factor of about 5.

TABLE 5.2 Comparison of Initiation Energies of Some Common Substances in Air and Pure Oxygen[32]
(Taken from Mackenzie, Ref. 32)

Flammable Substance	*Relative Amount of Energy Required to Initiate Combustion, millijoules*	
	In Air	*In Pure Oxygen*
Methane	0.3	0.003
Hydrogen	0.019	0.0012
Acetone	1.15	0.0024
Diethyl ether	0.2	0.0013

If no other data are available for determination of hazards, minimum ignition energies at ambient temperatures and pressures should be considered as approximately:

- 0.1 mJ for vapors in air
- 1.0 mJ for mists in air
- 10.0 mJ for dusts in air

STATIC ELECTRICITY

Introduction

Many apparently mysterious fires and explosions have eventually been traced to static. In spite of the large amount of information

about static electricity, it remains a complex phenomenon not often understood and appreciated. Static electricity is a potential source of ignition whenever there is a flammable mixture of gas or dust.

When two different or similar materials are in contact, electrons can move from one material across the boundary and associate with the other. If the two materials in contact are good conductors of electricity and are separated, the excess electrons in one material will return to the other before final contact is broken. But if one or both of the materials are insulators, this flow will be impeded. If the separation is done rapidly enough, some excess electrons will be trapped in one of the materials. Then both materials are "charged." Electric charges can build up on a nonconducting surface until the dielectric strength is exceeded and a spark occurs. The residual charge could ignite flammable mixtures.

The two materials or phases in initial contact may be:

- A single liquid dispersed into drops
- Two solids
- Two immiscible liquids
- A solid and a liquid
- A solid and a vapor or gas
- A liquid and a vapor or gas

The important thing to keep in mind is that whenever there is contact and separation of phases, a charge may be developed that could be disastrous. Three conditions must be met before an explosion caused by static electricity can take place:

1. An explosive mixture must be present.
2. An electric field must have been produced due to the electrostatic charge that had been generated and accumulated in a liquid or solid.
3. An electric field must be large enough to cause a spark of sufficient energy to ignite the mixture.

In designing preventive measures, all three factors should be controlled.

Static electricity is essentially a phenomenon of low current but high voltage and high resistance to current flow. A low-conductivity liquid flowing through a pipeline can generate a charge at a rate of 10^{-9} to 10^{-6} ampere (A). A powder coming out of a grinding mill can carry a charge at a rate of 10^{-8} to 10^{-4} A. At a charging rate of 10^{-6} A, the potential of a container insulated from earth can rise at a rate of 1000 volts/sec, and a voltage of 10,000 volts or higher can readily be obtained in this way.

Several electrostatic voltages and energies commonly encountered are typified by the following examples:

1. A person walking on dry carpet or sliding across an automobile seat can generate up to 5000 volts in dry weather. An individual having a capacitance of 100 pF (picofarads), a reasonable figure, could generate a spark energy of 1.25 mJ. This is far more than is needed to ignite some flammable vapor–air mixtures.
2. A person can accumulate dangerous charges up to about 20,000 volts when humidity is low.
3. A truck or an automobile traveling over pavement in dry weather can generate up to about 10,000 volts.
4. Nonconductive belts running over pulleys generate up to 30,000 volts. The voltage generated by a conveyor belt can be as high as 10^6 volts; the system can in effect act as a Van der Graaf generator.
5. The energy in the spark from an ordinary spark plug is 20 to 30 mJ.

The capacitance and the energy for ignition of people and of common objects are important. The capacitance of a human being is sufficient to ignite various flammable gas mixtures at commonly attained static voltages.

Hazard Determinants

Capacitance. The capacitance of an object is the ratio of the charge of the object to its potential. The capacitance gets larger as the object gets larger. With a given charge, the

voltage gets higher as the capacity of the object gets smaller. For a sphere, capacitance is given by:

$$C = Q(10^{-3})/V$$

The energy stored in a capacitor is:[27]

$$W = 0.5CV^2(10^{-3}) = 500Q^2/C$$

where:

C = capacitance, pF (1 pF = picofarad = 10^{-12} farad)
Q = charge, microcoulombs (1 coulomb = 1A/sec = charge on 6.2×10^{18} electrons)
V = voltage in kilovolts
W = energy, millijoules (mJ)

This energy may be released as a spark when the voltage gets high enough. The minimum sparking potential for charged electrodes is about 350 volts and occurs at a spacing of 0.01 mm. Sparks from an equally charged nonconductor are less energetic and may contain only part of the stored energy. These comparatively weak sparks are not likely to ignite dust clouds but can ignite flammable gases.

The energy that can be stored by capacitance of an object can be compared to the minimum ignition energies of flammable gas–air mixtures and of dust–air mixtures to determine the probability that a spark discharge may have sufficient energy to cause ignition. If the charged object is a poor conductor, the calculation of energy available to produce a spark may not be possible because the charge often is not uniformly distributed, and the resistance to flow of current is high. Figure 5.8 shows some typical values of electrical capacitance.[33]

Relaxation Time. When a liquid is flowing in closed metal pipes, static electricity is not a hazard. When the liquid enters a tank, it may become a hazard. Charges caused by liquid separation during pumping, flow, filtration, and other effects such as splashing and agitation can accumulate on the surface of the liquid in the tank and cause sparking

Fig. 5.8. Some typical values of electrical capacitance. (*Data from Eichel 1967.*[33])

between the liquid surface and the tank or conducting objects in the tank. The charge thus generated can be dissipated by relaxation or via discharge through a spark or corona discharge. The relaxation time is the time required for 63 percent of the charge to leak away from a charged liquid through a grounded conductive container. The half-time value is the time required for the free charge to decay to one-half of its initial value. Half-time is related to relaxation time by the relationship:

$$T_h = T_r \times 0.693$$

where:

T_h = half-time
T_r = relaxation time

Relaxation times vary from small fractions of a second up to minutes and even hours for some highly purified hydrocarbons that have very low conductivity.

It is important to recognize that a large charge can accumulate in the liquid even in a grounded container. In fact, it was reported that the majority of accidents attributed to static electricity in the petroleum industry have been with liquid in grounded containers.[2]

Relaxation time can be calculated as follows:

$$T_r = E(E_0/k)$$

where:

T_r = the relaxation time, in seconds; the time for 63% of the charge to leak away
E = relative dielectric constant, dimensionless
E_0 = absolute dielectric constant in a vacuum, $= 8.85 \times 10^{-14}$ to less than 1×10^{-18}
K = liquid conductivity, Siemens/cm

Siemens (S) are also called mhos.

Example: Benzene in a large tank could have a specific conductivity as low as 1×10^{-18} mho/cm and as high as 7.6×10^{-8} S/cm. The corresponding relaxation times for the two conductivities can be calculated as follows. Pure benzene has a dielectric constant of 2.5 to less than 1×10^{-18}. Using the above equation:

(1) $T_r = (8.85 \times 10^{-14}) \times (2.5)/(7.6 \times 10^{-8})$
$= 2.91 \times 10^{-6}$ sec

(2) $T_r = (8.85 \times 10^{-14}) \times (2.5)/(1 \times 10^{-18})$
$= 2.21 \times 10^{5}$ sec (This is in excess of 60 hours.)

Benzene typically contains some water and has a higher conductivity than in the above example and has a much lower relaxation time.

The purity of a liquid has a great effect on its relaxation time, and thus its static hazard potential. In actual practice, relaxation times of a few seconds to an hour are encountered, depending on purity and dryness of the liquid. This emphasizes the dangers of open sampling of tank contents soon after filling. If it is likely that the liquid being used has a low conductivity, it is important that enough time elapse between activities that can produce a static charge, such as loading a tank, and any activity that could cause a spark, such as sampling from the top of the tank.

In case (2) in the above example, a conductivity of 10^{-18} S/cm is so low that there may be little charge separation and little charge formation, and there may be no hazard even though the calculated relaxation time is extremely long. Materials with a half-time value of less than 0.012 sec have been reported not to cause a hazard. A useful rule to remember is that the concept of relaxation is very important because it is possible for liquid in a tank to retain an electric charge for a long time if the liquid is a poor conductor, even if the tank is grounded. The specific conductivity, and therefore the relaxation time, is greatly affected by impurities. For example, the specific conductivity of benzene can vary from as long as 1×10^{-18} to about 7.6×10^{-8} S/cm, depending on its purity. It can vary significantly with the amount of water or other materials dissolved in the benzene.[27]

Resistivity. The extent of charge separation is dependent on the resistivity of the liquid.

Some materials have a sufficiently high conductivity to render them safe in terms of static buildup. If the resistivity is low, charge separation is easy, but so is charge recombination through the liquid. If the resistivity is high, there may be appreciable charge separation without immediate recombination, leading to a high charge. If the resistivity is extremely high, there may be no charge separation, and there is no buildup of a charge. If the conductivity of a liquid falls in the hazardous range, it is possible to modify it by the use of a very small amount of an additive. Additives usually are a combination of a polyvalent metal salt of an acid such as carboxylic or sulfonic acid and a suitable electrolyte. Additives of this type can impart a conductivity of 10^{-8} S/m (Siemens per meter) in a 0.1 percent solution in benzene.[2]

A useful rule to remember is that when the resistivity of a liquid exceeds 10^{15} ohm centimeters, or is less than 10^{10} ohm centimeters, static generation or accumulation is negligible. Between these limits, the net generation of charges increases with the *maximum charge generation at 10^{13} ohm centimeters.* Styrene, for example, a commonly used monomer, has a resistivity of 4×10^{13} at 20°C,[34] and therefore is capable of building up a potentially hazardous charge.

Static Charge Development. Static electrification of solids can occur in various ways. Different operations will produce the percentages of the theoretical maximum charge density shown in Fig. 5.9.[33]

It should be noted that pure gases do not generate significant static electricity in transmission through pipes and ducts. Gases contaminated with rust particles or liquid droplets produce static, but this is not a problem in a closed, grounded piping system. *If these gases impinge on an ungrounded, conductive object, dangerous charges can accumulate on that object.* Wet steam, which contains water droplets, can develop charges. *If the water droplets contact an ungrounded conductor, that object can develop a static charge.*

Flammable gases may ignite when discharged to air during thunderstorms, even without a direct lightning hit. Dry hydrogen and occasionally other gases may ignite when they are discharged to air in normal weather. This may be so because the electric field

Fig. 5.9. Percentage of theoretical maximum charge produced by various operations. (*Data from Eichel 1967.*[33])

developed by the ejected gases can develop a corona discharge which can cause ignition. The minimum ignition energy of hydrogen is only 0.02 mJ. A toroidal ring developed by NASA (National Aeronautics and Space Administration) is reported to prevent unwanted discharge and subsequent ignition of a vent-stack outlet.[27]

Humidification. The conductivity of electrical nonconductors, such as plastics, paper, and concrete, depends on their moisture content. Relatively high moisture in these materials increases conductivity and therefore increases dissipation of static electricity. With relative humidity of 60 to 70 percent or higher, a microscopic film of moisture covers surfaces, making them more conductive.

Humidification can and often should be practiced to reduce the hazard of static electricity, but should not be relied on entirely to remove all possibility of static discharge.

In winter, cold air brought into a building and heated to normal room temperature is extremely dry, often less than 5 to 10 percent relative humidity. When processing solid materials that can develop a static charge, this air should be humidified to reduce static hazards as well as improve the comfort of personnel.

Filling Liquid Containers. A fire during top loading of a flammable liquid into a tank constitutes a serious problem if there can be a flammable mixture in the vapor space. Static electricity can be generated by splashing if the liquid is top-loaded, so it is normal practice to fill with a dip pipe positioned so the tip of the dip pipe is near the bottom of the tank. This may not be sufficient to prevent static charge buildup, as a charge may be generated in the bottom of the tank before the pipe tip is fully submerged, and it is possible for the liquid to acquire a charge before it reaches the tank.

Product filters using cotton, paper, felt, or plastic elements are prolific generators of static electricity. It is considered that at least 30 sec is necessary to dissipate this charge, although with dry nonconductive liquids, it may require as long as 500 sec.

Loading a less volatile liquid into a tank where there was previously a more volatile liquid is particularly hazardous because the more volatile liquid may form a flammable mixture, and the less flammable material is often a poor conductor and will not readily dissipate static charge. This type of loading accounts for 70 to 80 percent of severe losses at terminals.[2] This appears to occur most often when the compartments are one-fourth to one-third full, and when the temperature is close to $-1°C$.

Inerting the tank while it is being filled will reduce the possibility of ignition by static electricity and is highly recommended when it is possible and practical. However, this is not always practical. In any case, if inerts are to be used, they must be added carefully, as the following example illustrates. Two firemen were fatally injured when an explosion occurred as they were attempting to use portable CO_2 fire extinguishers to inert a tank truck. The source of ignition was believed to be a spark from the horn of the extinguisher to the latch on the tank truck. It was found that the voltage on the horn increased as the carbon dioxide "snow" passed down the horn to the outlet side.

Grounding and bonding lines, although very important, will not immediately dissipate the charge on the surface of a nonconducting liquid in a tank. A relaxation time for charge to be a dissipated should be allowed after filling or other operations to permit static charge on the liquid surface to dissipate to the dip pipe or tank shell. *The minimum time is one minute, but longer periods are advisable with some liquids that have extremely low conductivity. Bottom loading may reduce the static electricity hazard but does not eliminate it.*[35]

EXPLOSIONS

Development of Pressure

Exothermic reactions can lead to high temperatures and in the case of large fires to large

loss of property and severe damage from radiant energy. However, in many plant accidents it is the sudden generation of pressure that leads to severe damage, injury, and deaths. Hence, it can be stated that "pressure blows up plants, not temperature." Of course, temperature and pressure are closely related, but it is the pressure effect that is of concern in this section.

The word "deflagration" can be defined in several ways. One definition is "a reaction which propagates to the unreacted material at a speed less than the speed of sound in the unreacted substance."[27] Another definition of deflagration is from Latin, meaning "to burn down, or to burn rapidly with intense heat and sparks given off."[30] A deflagration may be an explosion, but not all deflagrations are explosions (a violently burning fire may be a deflagration, but that is not an explosion). On the other hand, not all explosions are deflagrations (a steam boiler may explode, but that is not a deflagration).

An explosion is a sudden and violent release of energy. Usually it is the result, not the cause, of a sudden release of gas under high pressure. The presence of a gas is not necessary for an explosion. An explosion may occur from a physical or mechanical change, as in the explosion of a steam boiler, or from a chemical reaction. The explosion of a flammable mixture in a process vessel may be either a deflagration or a detonation, which differ fundamentally. Both can be very destructive. Detonations are particularly destructive, but are unlikely to occur in vessels.

A detonation is a reaction that propagates to unreacted material at a speed greater than the speed of sound in the unreacted material; it is accompanied by a shock wave and extremely high pressures for a very short time. It is debatable whether the flammable range is the same as the detonable range. Detonation limits normally are reported to be within the flammable limits, but the view is widely held that separate detonation limits do not exist.

Unconfined vapor clouds can both deflagrate and detonate, with a deflagration being much more likely. A detonation is more destructive, but a deflagration also can produce a damaging pressure wave. A deflagration can undergo transition to a detonation in a pipeline, but this is most likely in vessels.

If a flammable mixture may be present in process equipment, precautions should be taken to eliminate ignition sources. However, it is prudent to assume that, despite these efforts, a source of ignition will at some time occur.

Deflagration

The conditions for a deflagration to occur are that the gas mixture is within the flammable range and that there is a source of ignition or that the mixture is heated to its autoignition temperature.

For the burning of hydrocarbon–air mixtures:

$$\frac{P_{2\mathrm{MAX}}}{P_1} = \frac{N_2 T_2}{N_1 T_1} = \frac{M_1 T_2}{M_2 T_1}$$

where:

T = absolute temperature
M = molecular weight of gas mixture
N = number of moles in gas mixture
P = absolute pressure
1,2 = initial and final states
2MAX = final state maximum value

The maximum pressure rise for a deflagration of flammable mixtures is approximately as follows for initial absolute pressures of 1 to 40 bar, for initial temperatures of 0 to 300°C, and for relatively small volumes of a few cubic meters:

$$\frac{P_2}{P_1} = \text{approx. 8 for hydrocarbon–air mixtures}$$

$$\frac{P_2}{P_1} = \text{approx. 16 for hydrocarbon–oxygen mixtures}$$

For conventionally designed pressure vessels:

$$\frac{P_b}{P_1} = \text{approx. 4 to 5}$$

where:

P_b = vessel bursting pressure
P_1 = normal design pressure
P_2 = pressure caused by deflagration

Therefore, in the absence of explosion relief, the deflagration explosion of a hydrocarbon–air mixture is easily capable of bursting a vessel if it is operating near its design pressure when the deflagration takes place. For reactions operating at or near atmospheric pressure, such as many drying and solids processing operations, it may be practical to construct facilities that will withstand the maximum explosion pressure of most dust–air and flammable gas–air mixtures.

Detonations

Detonation of a gas–air mixture may occur by direct initiation of detonation by a powerful ignition source or by transition from deflagration. This transition occurs in pipelines but is most unlikely in vessels. Two useful rules are:

1. Almost any gas mixture that is flammable is detonable if initiated with a sufficiently energetic source.
2. Detonation of a gas–air mixture is possible in pipelines but is unlikely in vessels.

Bartknecht[8] states that the range of detonability is narrower than the range of flammability. For example, the range of detonability of hydrogen in air is 18 to 59 vol. %, compared with the flammability of 4 to 75 vol. %. With flammable gases in air, if the length-to-diameter ratio of a pipe or vessel is more than about 10:1, and the pipe diameter is above a critical diameter, 12 to 25 mm, a detonation is possible.

Detonation Pressure. In the case of the burning of a flammable mixture of gases in a pipe with one end closed, a series of pressure waves traveling at the speed of sound moves through the unburned gas. Later waves traveling through the unburned gas, which has been heated by compression from the earlier waves, speed up because of the higher temperature and overtake the first wave, and a shock wave develops. Flame follows the shock wave and catches up with it, forming a detonation wave. A stable detonation wave may develop, which moves with supersonic speed relative to the unburned mixture, and peak incident (side-on) pressures are of the order of *30 times* the initial absolute pressure.

TABLE 5.3 Overpressure from Detonations[36]

	Pressure MPa	*Pressure lb/in²*
Incident overpressure	3.5	510
Maximum reflected pressure (wave strikes surface head-on)	28	4100
Load the structure feels (due to acceleration)	56	8100

(MPa means pressure in megapascals.)

Reflected Pressure. Reflected pressure increases the pressure on a rigid surface if the shock wave impinges on the surface at an angle to the direction of the propagation of the wave. The maximum ratio of reflected pressure to incident (side-on) pressure when a strong shock wave strikes a flat surface head-on is 8:1. Furthermore, acceleration from a suddenly applied force of the detonation wave can double the load that a structure "feels." Table 5.3 shows overpressure that can be expected from typical detonations.[36]

Thus, the stable detonation wave may cause enormously high pressures at closed ends of pipes, bends, and tees, where the greatest destruction from a gaseous detonation may occur.

Geometry. The following are some factors to consider when detonation is possible:

- Large length-to-diameter ratios promote the development of detonations; vessels should be designed with the lowest length-to-diameter ratio practicable if a detonation is possible.
- Equipment such as tanks (not including pipelines) designed to withstand 3.5 MPa (about 500 psig) usually will be adequate to contain a detonation, with other safeguards, for flammable gases in air at atmospheric pressure.

- Dished heads survive detonations better than do flat heads because of the more unfavorable incidence of flat heads.
- If turns in a process line are necessary, two 45° bends or a long sweep elbow will greatly reduce reflected pressure compared to a single 90° elbow.
- Restrictions such as orifices in pipelines may intensify a detonation by promoting pressure piling, which results when there are interconnected spaces such that the pressure rise in one space causes a pressure rise in a connected space. The enhanced pressure in the latter then becomes the starting pressure for a further explosion.
- Detonation may be extinguished when it enters a wider pipe from a smaller one, but the detonation may be regenerated somewhere along the longer pipe.
- Flame arrestors, if properly designed, can arrest detonations.

Explosion Violence

The *rate of pressure rise* is a measure of the violence of an explosion. The maximum rate of pressure rise for confined explosions is greatly affected by the volume of the vessel, with the influence of vessel volume on the rate of pressure rise being given by the following equation:

$$(dp/dt)_{\max}(V^{1/3}) = \text{a constant} = K_G$$

where:

$(dp/dt)_{\max}$ = maximum rate of pressure rise, bar/sec
V = vessel volume, cubic meters
K_G = a specific material constant, $(\text{bar})(\text{m})(\text{sec})^{-1}$

This is the *cubic law*, which states that for a given flammable gas, the product of the maximum pressure rise and the cube root of the vessel volume is a specific material constant, K_G.

The cubic law allows the prediction of the course of an explosion of a flammable gas or vapor in a large vessel, based on laboratory tests. It is valid only for the following conditions:

1. The same optimum concentration of the gas–air mixture.
2. Same shape of reaction vessel.
3. The same degree of turbulence.
4. The same ignition source.

Thus, to characterize an explosion, it is not enough to quote the maximum rate of pressure rise: the volume, vessel geometry, turbulence, and ignition energy must also be stated. Table 5.4 lists the K_G values for some common flammable gases measured under laboratory conditions.

It can be seen that the violence of an explosion with propane is about 1.5 times higher than one with methane, and one with hydrogen is about 10 times higher than one with methane. The explosive behavior of propane is representative of many flammable organic vapors in air. Some important relationships among pressure, temperature, turbulence, and vessel shape are discussed below.

1. *Explosion pressure is primarily the result of temperature reached during combustion, not a change in moles.* With complete combustion of propane in air there is a negligible change in moles of gas:

$$C_3H_8 + 5O_2 + \overbrace{18.8N_2}^{\{\text{----air----}\}} = 3CO_2 + 4H_2O + 18.8N_2$$

Number of moles at start = 24.8.
Number of moles after complete combustion = 25.8.

TABLE 5.4 K_G Values of Gases, Spark-Ignited with Zero Turbulence, Ignition Energy ~10 Joules, $P_{\max}$ = 7.4 bar[8]
(From Bartknecht, *Explosion Course, Prevention Protection*, Ref. 8, by Permission.)

Flammable Gas	K_G *(bar)(meter)/sec*
Methane	55
Propane	75
Hydrogen	550

Therefore, explosion pressure usually develops principally from an increase in temperature, not an increases in gas moles, during the combustion process of many materials.

Peak explosion pressure at constant volume occurs near the stoichiometric concentration in air. If only a small part of the total volume of a container is filled by an explosive gas–air mixture at atmospheric pressure, and the remainder of the vessel contains air, an explosion of this mixture can create enough pressure to severely damage containers that are designed to withstand only slight pressure —such as buildings and low-pressure storage tanks.

2. *Initial pressure affects maximum explosion pressure and rate of pressure rise.* If the initial pressure is increased above atmospheric pressure, there will be a proportional increase in the maximum explosion pressure and in the rate of pressure rise. Reducing the initial pressure will cause a corresponding decrease in maximum explosion pressure until finally an explosion reaction can no longer be propagated through the gas mixture.

3. *Initial temperature affects maximum explosion pressure and rate of pressure rise.* The maximum explosion pressure decreases when the starting temperature increases at the same starting pressure because of the lower density and thus smaller mass of material within a confined volume at higher temperatures. The maximum rate of pressure rise, $(dp/dt)_{max}$, increases as the initial temperature rises because the burning velocity increases with an increase in initial temperature.

4. *Initial turbulence increases rate of pressure rise.* Initial turbulence greatly increases the rates of explosion-pressure rise.[8,27] It has been found that with pentane and methane mixtures in air, $(dp/dt)_{max}$ can be five to nine times more with high initial turbulence than with no turbulence. The maximum explosion pressure is raised by about 20 percent. The course of explosions of flammable gases with a low normal speed of combustion, such as methane, is influenced by turbulence to a much higher degree than is the course of explosions with a high speed of combustion, such as hydrogen. Test data usually are obtained in equipment with a high degree of turbulence.

5. *Effect of vessel shape and increased initial pressure.* The maximum explosion pressure in confined vessels is not significantly affected by the volume or shape of the vessel in confined explosions for vessels that approximate the "cubic shape," that is, with a ratio of diameter to length (or vice versa) of about 1:1 to 1:1.5. In closed elongated vessels with central ignition, spherical ignition of the flame front will cause the flame to proceed swiftly in an axial direction. In the process, it compresses the unburned gases ahead of it, causing the violence of the explosion to increase, and pressure oscillations may occur.

DAMAGE ESTIMATES[37]

Damage estimates deal with the consequences of explosions and thermal radiation to both people and property. Physical models for explosions and thermal radiation generate a variety of incident outcomes: shock wave overpressure estimates, fragment velocities, and radiant flux. These models rely on the general principle that severity of outcome is a function of distance from the source of release. In addition to estimating the damage resulting from an explosion, it is also necessary to estimate how the consequences of these incident outcomes depend on the object of the study. To assess effects on human beings, damage estimates may be expressed as deaths or injuries. If physical property is the object, the damage estimates may be expressed as monetary losses.

Explosion Consequences

A principal parameter characterizing an explosion is the overpressure. Explosion effect modeling generally is based on TNT explosions to calculate the overpressure as a function of distance. Although the effect of a TNT explosion differs from that of a physical or a chemical explosion (particularly in the near-field), the TNT model is the most popular because a large data base exists for

TNT explosions. Several kinds of energy may be released in an explosion; three basic types are: (1) physical energy, (2) chemical energy, and (3) nuclear energy. Nuclear energy is not considered here. Physical energy may take such forms as pressure energy in gases, strain energy in metals, or electrical energy. Chemical energy derives from a chemical reaction. Examples of explosions involving chemical energy are runaway exothermic reactions, including decomposition and polymerization.

Table 5.5 summarizes the effects of explosion overpressure on structures. With respect to human casualties, heavy building damage usually is equated to a fatal effect, as the people inside the buildings probably would be crushed. People outside of buildings or structures are susceptible to direct blast injury (blast overpressure) and indirect blast injury (missiles or whole body transition).

Relatively high blast overpressures (>15 psig) are necessary to produce a human fatality from a direct blast. Instead, the major threat is produced by missiles or by whole body translation. Fatalities arising from whole body translation are mainly due to head injury from decelerative impact. Injury to people due to fragments usually results from either penetration by small fragments or blunt trauma from large fragments. TNO[39] suggested that projectiles with a kinetic energy of 100 Joules can cause fatalities. Table 5.6 shows damage to people (physiological damage) as a function of overpressure.

TABLE 5.5 Effect of Explosion Overpressure on Structures
(Courtesy of Center for Chemical Process Safety, American Institute of Chemical Engineers[38]

Pressure psi	*Damage*
0.02	Annoying noise (137 dB if of low, 10–15 Hz frequency).
0.03	Breaking of large glass windows under strain.
0.04	Loud noise (143 dB), sonic boom, glass failure.
0.10	Breakage of small glass windows under strain.
0.15	Typical pressure for glass breakage.
0.30	"Safe distance" (probability 0.95 of no serious damage below this value); projectile limits; some damage to house ceilings; 10% window glass broken.
0.40	Limited minor structural damage.
0.5–1.0	Large and small windows usually shattered; occasional damage to window frames.
0.70	Minor damage to house structures.
1.00	Partial demolition of houses; made uninhabitable.
1–2.00	Corrugated asbestos shattered; corrugated steel or aluminum panels, fastenings fail, followed by buckling; wood panel fastenings of standard housing fail; panels blown in.
1.30	Steel frames of clad buildings slightly distorted.
2.00	Partial collapse of walls and roofs of houses.
2–3.00	Concrete or cinder blocks shattered if not reinforced.
2.30	Lower limit of serious structural damage.
2.50	50% destruction of brickwork of houses.
3.00	Heavy machines (300 lbs), industrial buildings suffered little damage; steel frame buildings distorted and pulled away from foundation.
3–4.00	Frameless, self-framing steel panel building demolished; rupture of oil storage tanks.
4.00	Cladding of light industrial buildings ruptured.
5.00	Wooden utility poles snapped.
5–7.00	Nearly complete destruction of houses.
7.00	Loaded railcars overturned.
7–8.00	Brick panels, 8–12 in. thick, not reinforced, fail by shearing or flexure.
9.00	Loaded train boxcars completely demolished.
10.00	Probable total destruction of buildings; heavy machine tools (7,000 lb) moved and badly damaged; very heavy machine tools (12,000 lb) survive.
300.00	Limit of crater lip.

TABLE 5.6 Physiological Damage as a Result of Overpressure

Effect	*Peak Overpressure, psi*
Knock down	1.0
Ear drum damage	5.0
Lung damage	15
Threshold for fatalities	35
50% fatalities	50
99% fatalities	65

Radiation Consequences

The effect of thermal radiation on people and objects is determined by one of two approaches:

1. Simple tabulations based on experimental results.
2. Theoretical models based on the physiology of the skin burn response.

Data on time to pain threshold[40] are summarized in Table 5.7. For comparison, solar radiation intensity on a clear, hot summer day is about 320 Btu/hr ft^2 (1 kW/m^2). Other criteria for thermal radiation damage are shown in Table 5.8.[38]

The effect of thermal radiation on structures depends on whether they are combustible or not, and the nature and duration of the exposure. Thus, wooden materials will fail because of combustion, whereas steel will fail because of thermal lowering of the yield stress.

TABLE 5.7 Time to Pain Threshold for Varying Levels of Radiation[40]
(*Courtesy American Petroleum Institute*)

Radiation Intensity (Btu/hr ft^2)	*Radiation Intensity (kW/m^2)*	*Time to Pain Threshold (s)*
500	1.74	60
740	2.33	40
920	2.90	30
1500	4.73	16
2200	6.94	9
3000	9.46	6
3700	11.67	4
6300	19.87	2

TABLE 5.8 Effects of Thermal Radiation
(Courtesy of Center for Chemical Process Safety, American Institute of Chemical Engineers[38]

Radiation Intensity (kW/m^2)	*Observed Effect*
37.5	Sufficient to cause damage to process equipment.
25.0	Minimum energy required to ignite wood at indefinitely long exposures.
12.5	Minimum energy required for piloted ignition of wood, melting of plastic tubing.
9.5	Pain threshold reached after 6 seconds; second-degree burns after 20 seconds.
4.0	Sufficient to cause pain to personnel if unable to reach cover within 20 seconds; however, blistering of the skin (second degree burns) is likely; 0% lethality.
1.6	Will cause no discomfort for long exposure.

Unconfined Vapor Cloud Explosions (UVCE)

When a large amount of volatile material is released rapidly to the atmosphere, a vapor cloud forms and disperses. If the cloud is ignited before it is diluted below its lower flammability limit, an uncontrolled vapor cloud explosion will occur. This is one of the most serious hazards in the process industries. Both shock waves and thermal radiation will result from the explosion, with the shock waves usually the more important damage producers. UVCEs usually are modeled by using the TNT model.[38] The energy of the blast wave generally is only a small fraction of the energy available from the combustion of all the material that constitutes the cloud; the ratio of the actual energy released to that available frequently is referred to as the "explosion efficiency." Therefore, the TNT weight equivalent of a UVCE includes an explosion efficiency term, which typically is an empirical factor ranging from 1 to 10 percent. The explosion effects of a TNT charge are well documented.

Physical Explosions

A physical explosion usually results from the production of large volumes of gases by nonchemical means. The gases necessary for a physical explosion may be those already existing, such as compressed nitrogen released suddenly from a ruptured cylinder, or steam released explosively from a crack in a steam drum.

The following are some settings and situations in which physical explosions have been known to take place:

- Steam boilers.
- Hydraulic overfill of tanks or pipes with external applied pressure (as in pressure testing).
- Compressed air tanks.
- Deadheaded pumps.
- Thermal expansion of tanks or pipes.
- Liquid cryogenic fluids on water (such as liquid methane on water).
- Water suddenly mixed with sulfuric acid (also may cause a chemical explosion).
- BLEVE (Boiling Liquid Expanding Vapor Explosion) with superheated liquid (flammable or nonflammable) (see next section).
- Explosion of grinding wheel at too high a speed.
- Liquid water in molten $MgCl_2$ solution at high temperatures.
- Implosions due to vacuum.
- Overpressured refrigerant systems.
- Molten metals exploding violently on contact with water.
- Some molten metals exploding when mixed with each other.
- The mixing of two immiscible liquids whose boiling points are not widely separated.

Steam boilers are commonly used in power plants and industries of all kinds. They generally are taken for granted now, but in the second half of the nineteenth century boilers blew up with alarming regularity. Records indicate that from 1870 to 1910 there were at least 10,000 boiler explosions in the United States and adjacent areas of Canada and Mexico; that is, more than one recorded explosion every 36 hours! By 1910, the rate had jumped to between 1300 and 1400 per year. On October 8, 1894, in the Henry Clay Mine in Shamokin, Pennsylvania, 27 boilers disintegrated almost simultaneously! Mainly because of the incorporation of the ASME Boiler Code into laws, boiler explosions have decreased dramatically.[41]

When a pressurized vessel ruptures, the resulting stored energy is released. This energy can cause a shock wave and accelerate vessel fragments. If the contents are flammable, ignition of the released gas could produce additional effects. There is a maximum amount of energy in a bursting vessel that can be released, and it is released in the following proportions:[36]

	Distribution of Energy When Vessel Ruptures		
Type of Failure	*Strain Energy*	*Kinetic Energy of Fragments*	*Shock Wave Energy*
Brittle Failure	<10%	~20%	up to 80%
Plug Ejection	small	up to 60 to 80%	20 to 40%

The relative distribution of these energy components will change over the course of the explosion, but most of the energy is carried by the shock wave with the remainder going to fragment kinetic energy. To estimate the damage resulting from the shock wave from a physical explosion, the TNT model is used widely. To determine the TNT equivalent of a physical explosion, the total energy in the system must be estimated. For a physical explosion, if the expansion occurs isothermally, and ideal gas laws apply, then the TNT equivalent of the explosion can be calculated. This energy then can be used to estimate overpressure at any distance from the explosion. The analogy of the explosion of a container of pressurized gas to a point source explosion of TNT is not appropriate in the near-field. Prugh[42] suggests a correction method using a virtual distance R_v from an explosion center.

In addition to shock wave effects, a major hazard of a ruptured gas-filled vessel is from

projectiles. To estimate damage from projectiles, both the initial velocity and the range are required. A simplified method for calculating the initial velocity uses the following equation:[43]

$$u = 2.05(PD^3/W)^{0.5}$$

where:

u = initial velocity, ft/sec
P = rupture pressure psig
D = fragment diameter, inches
W = weight of fragments, pounds

Clancey[44] gives the following values for initial velocity for the majority of fragments from a TNT explosion:

- Thin case: 8000 ft/sec
- Medium case: 6000 ft/sec
- Thick case: 4000 ft/sec

Once the initial velocity has been determined, the maximum range of the fragment, ignoring air resistance, can be estimated from:

$$R_{\max} = \frac{u^2}{g}$$

If the above values for typical velocities are substituted into the above equation, a maximum range of 5×10^5 feet is possible. Therefore, it is clearly necessary to include air resistance. To include air resistance, a value of C_D, the drag coefficient, must be estimated. The drag coefficient ranges from 0.48 for a sphere to 2 for flow perpendicular to a flat strip, and for most fragments ranges from 1.5 to 2.0.

If one knows the air density, drag coefficient, exposed area of the fragment, mass of the fragment, and the initial velocity, the maximum range R can be calculated with the aid of Fig. 5.10.[45] Although this technique gives the maximum range, most fragments do not travel the maximum distance but fall at distances between 0.3 and 0.8 of the maximum.

The energy required to impart an initial velocity of u to a fragment is:

$$E = \tfrac{1}{2}mu^2$$

where:

m is the mass of the fragment (lb)
u the initial velocity (ft/sec)

Example. A high pressure vessel containing air at 600 bar has ruptured, leading to 15 fragments of approximately equal mass (85 lb), one of which was found as far as 400 feet from the vessel. This fragment has a drag coefficient of 1.5 and an exposed area of 3 ft.^{2} Assuming that 20 percent of the explosion energy went to energy of the fragments, estimate the energy of the explosion in weight equivalent TNT. The air density is 0.081 lb/ft.^{3}

Procedure. Assuming that the fragment found at 400 feet is at the maximum range for the fragments, the scaled fragment range R_s can be calculated:

$$R_s = \frac{r_0 C_D A_D R}{M}$$

$$= \frac{0.081 \times 1.5 \times 3 \times 400}{85}$$

$$= 1.7$$

From Fig. 5.10, we obtain a scaled force (F_s) of approximately 5. The initial velocity of the fragment then can be calculated as:

$$u = \sqrt{\frac{MgF_s}{r_0 C_D A_D}}$$

$$= \sqrt{\frac{85 \times 32.17 \times 5}{0.081 \times 1.5 \times 3}}$$

$$= 194 \text{ ft/sec}$$

The energy required to give the fragment this initial velocity is:

$$E = \tfrac{1}{2}(85)(194)^2$$

$$= 1.6 \times 106 \text{ lb ft}^2/\text{sec}^2$$

$$= 64 \text{ Btu}$$

Since there were 15 fragments, the total energy of the explosion that went into

where: $F_s = \frac{r_o C_D A_D u^2}{Mg}$ $R_s = \frac{r_o C_D A_D R}{M}$

and: r_o = density of air lb/cu ft C_D = drag coefficient

A_D = exposed area of fragment, ft2 R = max range of fragment, ft

M = mass of fragment, lb U = fragment initial velocity, fps

$g = 32.17 \text{ ft. lb.}_m/\text{lb}_f\text{s}^2$

Fig. 5.10A. Scaled fragment range vs. scaled force. (*Baker et al. 1983.*[45])

Fig. 5.10B. Maximum horizontal range of blast fragments. (*Clancy*[44].)

fragment kinetic energy is 15 × 635.8 Btu = 9537 Btu. If only 20 percent of the explosion energy went into fragment kinetic energy, then the total explosion energy is 47,680 Btu, which is the equivalent of 23.8 lb of TNT. Using the method of Clancy[44], 2.4 lb TNT can provide a maximum range of 950 ft for projectiles (Fig. 5.10B).

BOILING LIQUID EXPANDING VAPOR EXPLOSIONS (BLEVEs)

Among the most damaging of accidents is a Boiling Liquid Expanding Vapor Explosion (BLEVE, pronounced BLEV-ee). This occurs when a pressure vessel containing liquid is heated so that the metal loses strength and ruptures. Typically, this happens when the vessel failure results from overheating upon exposure to fire. The failure usually is in the metal contacting the vapor phase; the metal in this area heats to a higher temperature because there is no liquid heat sink to keep the metal temperature from rising rapidly, as there is where metal contacts a liquid phase. A BLEVE can occur with both flammable materials and nonflammable materials, such as water. In all cases the initial explosion may generate a blast wave and missiles. If the material is flammable, it may cause a fire or may form a vapor cloud that then gives rise to a secondary explosion and fireball. Kletz states that BLEVEs can cause as many casualties as can unconfined vapor cloud explosions.[2]

The best known type of BLEVE involves LPG (liquefied petroleum gas). Once a fire impinges on the shell above the liquid level,

the vessel usually fails within 10 to 20 minutes. In the case of a BLEVE involving a flammable material, the major consequences are, in order of decreasing importance:

1. Thermal radiation from the resultant fireball.
2. Fragments produced when the vessel fails.
3. Blast wave produced by the expanding vapor/liquid.

For example, a BLEVE of a propane sphere with a diameter of 50 feet, holding about 630,000 gallons, could cause damage as far away as 13,600 feet, and radiation damage and fragmentation damage would each extend to about 3000 feet.

In a fire, a tank containing liquid is most vulnerable in the shell at the vapor space because very little heat can be absorbed by the vapor, and the metal in the vapor space can heat up rapidly to a temperature where it will weaken rapidly. The metal contacting the liquid phase will heat up much less rapidly because the liquid can absorb significant amounts of heat, keeping the shell temperature down in that area for a significant amount of time. Thus, there is a dilemma: a partly full vessel may BLEVE sooner than will a full vessel, but a full vessel will have more fuel for the resulting fireball and fire than will a partly empty vessel.

Significant equipment and building damage frrom radiation is possible from a BLEVE. Wooden structures may be ignited if the radiant heat density at the structure's location exceeds the threshold value for ignition of wood. Severe damage from fragmentation can be expected in the area where 50 percent or more of the fragments may fall, or typically about 300 feet from the vessel.

A BLEVE can lead to shock waves, projectiles, and thermal radiation. The effects of a shock wave and projectiles were dealt with earlier; by far the most serious consequence of a BLEVE is the radiation received from the fireball. The following calculational procedure is used to determine thermal impact (details are available in reference 38):

MECHANICAL HEAT

Mechanical motion in fluids becomes kinetic energy and may become heat in devices with rotating parts. Mechanical heat input from rotating agitators, pump impellers, and other mechanical equipment must be taken into account in the design of process equipment, particularly in systems containing reactive chemicals. This section will provide some guidelines for the analysis of individual cases involving pumps and agitated tanks.[46–48]

Some useful rules are as follows:

1. A deadheaded pump is a pump operating full of liquid and with inlet and outlet valves closed.
2. Almost all deadheaded centrifugal pumps with motors of three horsepower or larger are headed for trouble if left deadheaded. (Depending on the horsepower, a few minutes may be too long.)

3. The heat input from the rotating impeller in a deadheaded centrifugal pump is always a large value relative to the heat sink of the fluid and the pump.
4. It is not necessary for there to be a chemical reaction in a pump for an explosion to take place. Deadheaded pumps containing only water or brine have blown up.
5. An agitator or a circulating pump left on in a vessel of a reactive chemical may heat up the contents enough to cause a runaway reaction.
6. All centrifugal pumps with motors larger than three horsepower should be protected in some way to prevent deadheading. A temperature alarm in the casing is a minimum form of protection. A better way may be to have the high-temperature alarm wired to the process control computer, to both alarm and shut off the pump. Other systems are available and may be used; they may include (but are not limited to) a relief valve on the pump, a minimum flow valve, and a flow orifice in the recirculating line. A relief valve on a pump relieving back to the pump inlet may not eliminate the problem of heat buildup in a deadheaded pump and usually should be avoided unless other protective measures are used such as a high-temperature device.
7. An ammeter on the pump motor usually is not a reliable means of detecting deadheaded conditions. The low power factors often experienced with pump motors, and the nature of pump curves, often make it difficult to distinguish between normal running and deadheaded conditions using an ammeter.
8. For mechanical heat equivalent, the following is recommended: (a) For pumps, use 50 percent of the connected motor horsepower for centrifugal pumps that are deadheaded, unless better information is available. (b) For agitators, use 100 percent of the vendor rated shaft input horsepower for the input shaft (total power less drive and bearing inefficiencies) for the actual material in the vessel.

VACUUM[49]

Ask any chemical engineers who have had some plant experience what they know about vacuum, and they probably will smile and tell a tale about some piece of equipment that tried to turn itself inside out. Usually no one was hurt, and often there is no massive leakage—but not always!

The design for the internal pressure condition of vessels usually is straightforward and well understood. Under vacuum conditions, equipment is subject to external pressure from the atmosphere; and the design for external pressures is more difficult than that for internal pressures. The devious ways in which external pressure can be applied often may be overlooked.

The following are some obvious causes of vacuum collapse:

1. Liquid withdrawal by pump or gravity draining.
2. Removal of gas or vapor by withdrawing with a blower, fan, or jet.
3. Siphoning of liquids.

Less obvious causes include:

1. Condensation of vapor.
2. Cooling of hot gas.
3. Combination of cooling and condensation of a mixture of gas and condensable vapor.

Sometimes obscure causes of vacuum collapse include:

1. Absorption of a gas in a liquid; for example, ammonia in water, carbon dioxide in water, hydrogen chloride in water.
2. Reaction of two or more gases to make a liquid or solid; for example, ammonia plus hydrogen bromide to form ammonium bromide.
3. Reaction of a gas and a solid to form a solid; for example, corrosion in a tank, air plus Fe or FeO forming Fe_2O_3 in the presence of water.

4. Reaction of a gas and a liquid to give a liquid; for example, chlorination, hydrogenation, ethylation.
5. Sudden dropping of finely divided solids in a silo, creating a momentary vacuum that can suck in the sides of the silo.
6. Flame arrestors plugging; for example:
   (a) In styrene service, vapor may condense in flame arrestors, and the liquid formed is low in inhibitor; the liquid may polymerize and plug off the arrestor. Possible solutions: clean the arrestor frequently or use a PVRV (pressure-vacuum-relief valve).
   (b) Liquid service in cold weather: vapor may condense in a flame arrestor and the liquid formed may freeze and plug the arrestor. Possible solution: heat and insulate the arrestor to prevent condensation.
7. Maintenance and testing. It is not a good idea to apply vacuum on a vessel during maintenance or testing without full knowledge of the external pressure rating unless a suitable vacuum relief device is in place and operable.

Protective Measures for Equipment

If equipment may be subject to vacuum, consideration should be given to designing the equipment for full vacuum. This may eliminate the need for complicated devices such as vacuum relief valves and instruments; if they are used, designing the equipment for full vacuum will prevent collapse of the vessel if the instruments or relief valves fail or plug.

A disadvantage of this approach is that it usually is expensive. However, when the total cost of a suitably instrumented vessel not designed for vacuum is compared with the cost of a vessel designed for vacuum but without the extra equipment, the difference may be small or negligible, and the vessel designed for vacuum will be inherently safer. If a vessel is designed for vacuum, precautions should be taken to ensure that internal or external corrosion will not destroy the integrity of the vessel.

PROCESS SAFETY MANAGEMENT SYSTEMS

The public, customers, employees, and local and federal government regulatory agencies all demand that companies take the actions needed to reduce the possibility of the release of hazardous materials. From 1975 to 1991, a number of major chemical or chemical-related incidents had major impacts on the chemical industry and on many communities as well. As the chemical process industries have developed better ways to improve process safety, safety management systems have been introduced to help implement improvements in process safety engineering activities. Management systems for chemical process safety are comprehensive sets of policies, procedures, and practices designed to ensure that barriers to major incidents are in place, in use, and effective.[24,38] Following are some definitions that are useful in discussions on managing process safety.

Definitions

Accident. An unplanned event with undesirable consequences.

Consequences. Damage from a scenario; for example, the number of people exposed to a chemical cloud.

Event. The most elementary action in an accident; for example, an operator response or action.

Hazard. Result of combining a scenario and a consequence; for example, a chemical cloud is formed, and people are exposed. Also, the inherent potential of a material or activity to harm people, property, or the environment.

Incident. An occurrence that may be either an accident or a near miss.

Near miss. Extraordinary event that could have resulted in an accident but did not.

Probability. The likelihood of the occurrence of events or a measure of degree of belief, the values of which range from 0 to 1.

Quantitative Risk Analysis (QRA). The systematic development of numerical estimates of the expected frequency and/or consequence of potential accidents associated with a facility or an operation. Often used interchangeably with "quantitative risk assessment."

Risk. Result of combining scenario, consequence, and likelihood, for example, likelihood (probability) of cloud being formed and people being exposed, given the process specifics. Or, a measure of potential economic loss or human injury in terms of the probability of the loss or injury occurring and the magnitude of the loss or injury if it occurs. Or, combination of the expected frequency (events/year) and/or consequence (effects/event) of a single accident or a group of accidents.

Risk assessment. The systematic evaluation of the risk associated with potential accidents at complex facilities or operations. Or, the process by which the results of a risk analysis are used to make decisions, either through relative ranking of risk reduction strategies or through comparison with risk targets. The terms "risk analysis" and "risk assessment" often are used interchangeably in the literature.

Risk management. A part of chemical process safety management. Risk analysis (often used interchangeably with risk assessment) is a part of risk management.

Scenario. End result of a series of events, for example, release of a chemical cloud.

Some Tools for Evaluating Risks and Hazards

Dow Fire and Explosion Index. The Dow Fire and Explosion Index (F&EI), developed by The Dow Chemical Company, is an objective evaluation of the potential of a facility for a fire, an explosion, or a reactive chemical accident. Its purpose is to quantify damage from incidents, identify equipment that could contribute to an incident, and suggest ways to mitigate the incident; it also is a way to communicate to management the quantitative hazard potential. It is intended for facilities handling flammable, combustible, or reactive materials whether stored, handled, or processed. The goal of the F&EI evaluation is to become aware of the loss potential and identify ways to reduce the potential severity in a cost-effective manner. It does not address frequency (risk) except in a general way. The number is useful mainly for comparisons and for calculations of damage radius, maximum probable property damage, and business interruption loss, and to establish frequency of reviews. The method of carrying out an F&EI evaluation is available to the public from the American Institute of Chemical Engineers, 345 E. 47th St., New York, NY 10017.

Failure Modes and Effects Analysis (FMEA). FMEA is a systematic, tabular method for evaluating the causes and effects of component failures. It represents a "bottom-up" approach in contrast with fault tree, where the approach is "top-down." In large part, HAZOP is a well-developed form of FMEA.[2]

Fault Tree. A fault tree is a logical model that graphically portrays the combinations of failures that can lead to a particular main failure or accident of interest. A fault tree starts with a top event, which is usually a hazard of some kind. The possibility of the hazard must be foreseen before the fault tree can be constructed. A fault tree helps reveal the possible causes of the hazard, some of which may not have been foreseen.[2]

Safety Audit. A safety audit is a method of reviewing the actual construction and operation of a facility. Often safety audits are conducted by a small interdisciplinary team. At least some of the members of the team are not connected with the plant. The audit may be carried out before startup and also is repeated later at intervals of, typically, one to five years.

Chemical Exposure Index. The Chemical Exposure Index is a technique for estimating the relative toxic hazards of chemicals, developed by The Dow Chemical Company. It provides for the relative ranking of toxic chemicals in a given facility, including factors relating to toxicity, quantity volatilized, distance to an area of concern, and physical properties. A description of the method is found in *Guidelines for Safe Storage and Handling of High Toxic Hazard Materials,* Center for Chemical Process Safety, American Institute of Chemical Engineers.[50]

EVALUATION OF HAZARDS AND RISKS

HAZOP[2,50–53]

HAZOP stands for "Hazard and Operability Studies," a set of formal hazard identification and elimination procedures designed to identify hazards to people, process, plants, and the environment. The techniques aim to stimulate the imagination of designers and operations people in a systematic way so they can identify potential hazards. In effect, HAZOP studies assume that there will be mistakes, and provide a systematic search for these mistakes. In some European countries, HAZOP studies are mandatory and attended by observers from regulatory authorities to ensure that the studies are carried out correctly. The examination of accidents[52] during 1988 at a large U.S. chemical company revealed that the accidents could be classified as follows:

Spills: 52 percent
Emissions: 30 percent
Fires: 18 percent

Of the fires, about 50 percent occurred during construction, 25 percent were due to pump seal failure, and the remaining 25 percent resulted from engineering and operational oversights that a HAZOP study possibly could have prevented.

Of the emissions, 37 percent were due to piping failure, with lined pipe being the largest contributor. Operational and procedural issues accounted for 53 percent of the remainder.

Of the spills, 11 percent were due to equipment failures. Piping failures (especially lined pipe and gaskets) accounted for 30 percent, and 56 percent were caused by various types of operational errors, noncompliance with procedures, or nonexistent procedures.

Material handling was a factor in many spills and emissions.

The most frequent type of operational error was a valve being left in an improper position, either open or closed.

HAZOP studies probably could have reduced the number and the seriousness of the problems experienced.

Some investigations have shown that a HAZOP study will result in recommendations that are 40 percent safety-related and 60 percent operability-related. Thus, HAZOP is far more than a safety tool; a good HAZOP study also results in improved operability of the process or plant, which can mean greater profitability.

The HAZOP technique can be used to identify human error potential. From a practical point of view, human error and its consequences can occur at all levels of a management structure as well as in the operation of a particular plant or process. Carried out correctly, Technica[54] states that a HAZOP study will identify at least 70 to 75 percent of potential operational and safety problems associated with a particular design process, including human error.

The HAZOP technique also can be used for the evaluation of procedures. Procedures may be regarded as a "system" designed to "program" an operator to carry out a sequence of correct actions. Deviations from intent are developed, with the emphasis on "operator action deviation" rather than "physical property deviation." It is the procedure, not the hardware, that is the object of study, but hardware modifications may be recommended to cover potential problems identified from procedure deviations.

Operating Discipline

HAZOP can be an important part of establishing the operating discipline of a plant, which can be defined as the total information required to understand and operate the facility. By recording recommendations and actions in detail on all parts of the process, the rationale behind the way the process is designed and intended to operate and key details of the process will be available, which are rarely on record from other sources. This can be especially helpful when plant changes are made, and when new plants based on the same design are built.

Risk Analysis and Assessment

Risk is defined (see above) as the combination of expected frequency (events/year) and consequence (effects/event) of accidents that could occur as a result of an activity. Risk analysis evaluates the likelihood and consequence of an incident. Risk assessment is an extension of risk analysis that includes making judgments on the acceptability of the risk. It may be qualitative or quantitative, or it can range from simple, "broad brush" screening studies to detailed risk analyses studying large numbers of incidents and using highly sophisticated frequency and consequence models. Throughout the published literature, the terms "risk analysis" and "risk assessment" are used interchangeably.

Quantitative Risk Analysis (QRA)

Quantitative Risk Analysis (QRA) models the events, incidents, consequences, and risks, and produces numerical estimates of some or all of the frequencies, probabilities, consequences, or risks.[38,55] QRA can be done at a preliminary level or a detailed level, and in all cases may or may not quantify all events, incidents, consequences, or risks.[56] QRA is the art and science of developing and understanding numerical estimates of the risk associated with a facility or an operation. It uses highly sophisticated but approximate tools for acquiring risk understanding.

QRA can be used to investigate many types of risks associated with chemical process facilities, such as the risk of economic losses or the risk of exposure of members of the public to toxic vapors. In health and safety applications, the use of QRA can be classified into two categories:

1. Estimating the long-term risk to workers or the public from chronic exposure to potentially harmful substances or activities.
2. Estimating the risk to workers or the public from episodic events involving a one-time exposure, which may be acute, to potentially harmful substances or activities.

This discussion will focus on the use of QRA in the assessment of acute hazards and episodic events only.

QRA is fundamentally different from many other chemical engineering activities (e.g., chemistry, heat transfer), whose basic property data are capable of being theoretically and empirically determined and often established experimentally. Some of the basic "property data" used to calculate risk estimates are probabilistic variables with no fixed values, and some of the key elements of risk must be established by using these probabilistic variables. QRA is an approach for estimating the risk of chemical operations by using the probabilistic approach; it is a fundamentally different approach from those used in many other engineering activities because interpreting the results of QRA requires an increased sensitivity to uncertainties that arise primarily from the probabilistic character of the data.

The underlying basis of QRA is simple in concept. It offers methods to answer the following four questions:

1. What can go wrong?
2. What are its causes?
3. What are its consequences?
4. How likely is it?

It has been said that QRA attempts to answer the question, "Does this plant have

the potential to cause another Bhopal incident?" (This incident is described below.) Although the application of QRA might not have prevented the Bhopal incident, it is at least possible that the proper application of QRA could have made this incident extremely unlikely.

Methods of Showing Results of Risk Studies

F–N curves are used for societal risks; risk contours show individual risk. Figure 5.11 shows various contours illustrating the probability of a particular impact on an individual located on the contour line. Figures 5.12 and 5.13 illustrate two of the more commonly used methods for displaying societal risk results: an *F–N* curve and a risk profile. The *F–N* curve plots the cumulative frequencies of events causing *N* or more impacts, with the number of impacts (*N*) shown on the horizontal axis. With the *F–N* curve it is easy to see the expected frequency of accidents that could harm more than a specified number of people. The *F–N* curve is a cumulative illustration.

The risk profile shows the expected frequency of accidents of a particular category or level of consequence. The diagonal line is a line of constant risk defined such that the product of expected frequency and consequence is a constant at each point along the line. As the consequences of accidents go up, the expected frequency should go down in order for the risk to remain constant. If a portion of the histogram sticks its head above the line (i.e., if a particular type of accident contributes more than its fair share of the risk), then that risk is inconsistent with the risk presented by other accident types.

A method for graphically displaying individual risk results is use of the risk contour, or risk isopleth. If individual risk is defined as the likelihood of someone suffering a specified injury or loss, then individual risk can be calculated at particular geographic locations around the vicinity of a facility or

Fig. 5.12. Example of an *F–N* curve (societal risk).

Fig. 5.13. Risk profile—expected frequency of accidents of a particular category.

Fig. 5.11. Risk contour (individual risk). Probability per year of an average individual being severely injured.

an operation. If the individual risk is calculated at many points surrounding the facility, then points of equal risk can be connected to create a risk contour map showing the geographic distribution of the individual risk.

THE PRINCIPAL REASON FOR MOST CHEMICAL PROCESS ACCIDENTS

Ask any group of people experienced in chemical plant operations what causes most chemical process accidents, and you will get a variety of answers including: operator error, equipment failure, poor design, act of God, and bad luck. However, in the opinion of representatives of many of the large chemical and oil companies in the United States, these answers are generally incorrect. The Center for Chemical Process Safety, an organization sponsored by the American Institute of Chemical Engineers, includes representatives of many of the largest chemical and oil companies in the United States, and states that "It is an axiom that process safety incidents are the result of *management system failure*." Invariably, some aspect of a process safety management system can be found that, had it functioned properly, could have prevented an incident (or reduced the seriousness of it). "It is a rare situation where an 'Act of God' or other uncontrollable event is the sole cause of an incident. Much more common is the situation where an incident is the result of multiple causes, including management system failures. Therefore, it is more appropriate to presume that management system failures underlie every incident so that we may act to uncover such failures and then modify the appropriate management systems, rather than presume that if an 'Act of God' appears to be the immediate cause, investigation should cease because there is nothing that can be done to prevent such future incidents." [24]

For example, consider a case where a small amount of hazardous material is spilled while a sample is being taken from a process line. It is not enough to look into the situation and conclude that this is an example of an operator error where procedures were not followed, and then simply to recommend that the employee be instructed to follow procedures in the future. Further investigation may reveal deficiencies in the training system or in the equipment. Still more investigation may reveal deficiencies in the management system that plans resources for training or that provides for proper equipment for sampling. It then may be appropriate to change the management system to prevent repetition of the incident.

Levels of Causes

There are several levels of causes of accidents, usually (1) the immediate cause, (2) contributing causes to the accident or to the severity of the accident, and (3) the "root cause." The root cause is what really caused the accident, and when this is determined, it may be possibly to *prevent* future similar accidents. With the 20–20 hindsight that is available after an accident, the root cause usually can be found. The purpose of the discussion in the next section is to illustrate how knowledge about the root causes of some important accidents can help to keep them from happening again. It will be noted that the root cause is rarely the fault of one person, but instead is the result of a management system that does not function properly.

Following are brief analyses of several case histories that have been of landmark importance in the industrial world, and that have affected the chemical industry all over the world.

CASE HISTORIES

Flixborough, England, 1974[2]

On June 1, 1974, an accident occurred in the Nypro plant in Flixborough, England, in a process where cyclohexane was oxidized to cyclohexanone for the manufacture of caprolactam, the basic raw material for the production of Nylon 6. The process consisted of six reactors in series at 155°C and 8.8 bar (130 psig) containing a total of 120 tons of cyclohexane and a small amount of cyclohexanone. The final reactor in the process

contained 94 percent cyclohexane. There was a massive leak followed by a large unconfined vapor cloud explosion and fire that killed 26, injured 36 people, destroyed 1821 houses, and damaged 167 shops. It was estimated that 30 tons of cyclohexane was involved in the explosion. The accident occurred on Saturday; on a working day, casualties would have been much higher.

The accident happened when the plant had to replace one of six reactors and rushed to refit the plant to bypass the disabled reactor. Scaffolding was jerry-rigged to support a 20-inch pipe connecting reactor four with reactor six, which violated industry and the manufacturer's recommendations. The reactor that failed showed stress crack corrosion. The only drawings for the repair were in chalk on the machine shop floor. Both ends of the 20-inch pipe had expansion joints where they attached to the reactors. The pipe was supported on scaffolding-type supports and was offset with a "dog-leg" to fit the reactors, which were at different levels to promote gravity flow. The safety reviews, if any, were insufficient.

Immediate Cause. A pipe replacing a failed reactor failed, releasing large quantities of hot cyclohexane forming a vapor cloud that ignited.

Contributing Causes to the Accident and the Severity of the Accident:

1. The reactor failed without an adequate check on why (metallurgical failure).
2. The pipe was connected without an adequate check on its strength and on inadequate supports.
3. Expansion joints (bellows) were used on each end of pipe in a "dog-leg" without adequate support, contrary to the recommendations of the manufacturer.
4. There was a large inventory of hot cyclohexane under pressure.
5. The accident occurred during start-up.
6. The control room was not built with adequate strength, and was poorly sited.
7. The previous works engineer had left and had not been replaced. According to the Flixborough Report, "There was no mechanical engineer on site of sufficient qualification, status or authority to deal with complex and novel engineering problems and insist on necessary measures being taken."
8. The plant did not have a sufficient complement of experienced people, and individuals tended to be overworked and liable to error.

Root Cause. Management systems deficiencies resulted in:

1. A lack of experienced and qualified people.
2. Inadequate procedures involving plant modifications.
3. Regulations on pressure vessels that dealt mainly with steam and air and did not adequately address hazardous materials.
4. A process with a very large amount of hot hydrocarbons under pressure and well above its flash point installed in an area that could expose many people to a severe hazard.

This accident resulted in significant changes in England and the rest of the world in the manner in which chemical process safety is managed by industry and government. One of the conclusions reached as a result of this accident, which has had wide effect in the chemical industry, is that "limitations of inventory (or flammable materials) should be taken as specific design objectives in major hazard installations."

The use of expansion joints (bellows, in this case) which were improperly installed may have been a principal reason for the accident. This provides additional reasons not to use expansion joints (except in special exceptional circumstances).

Bhopal, 1985 (C&EN Feb. 11, 1985; Technica 1989[54])

On December 3 and 4, 1985, a chemical release causing a massive toxic gas cloud occurred at the Union Carbide India, Ltd., plant in

Bhopal, India. The process involved used methyl isocyanate (MIC), an extremely toxic chemical, to make Sevin, a pesticide. According to various authoritative reports, about 1700 to 2700 (possibly more) people were killed, 50,000 people were affected seriously, and 1,000,000 people were affected in some way. The final settlement may involve billions of dollars. It was one of the worst industrial accidents in history. The accident occurred when about 120 to 240 gallons of water were allowed to contaminate an MIC storage tank. The MIC hydrolyzed, causing heat and pressure, which in turn caused the tank rupture disk to burst.

Equipment designed to handle an MIC release included a recirculating caustic soda scrubber tower and a flare system designed for 10,000 lb/hr, which would be moderate flows from process vents. It was not designed to handle runaway reactions from storage. The design was based on the assumption that full cooling would be provided by the refrigeration system. The actual release was estimated to be 27,000 lb over 2 hours, with the tank at 43°C. At the time of the release the refrigeration had been turned off. The flare tower was shut down for repairs. A system of pressurized sprinklers that was supposed to form a water curtain over the escaping gases was deficient, in that water pressure was too low for water to reach the height of the escaping gas.

There have been conflicting stories of how the water got into the tank, including operator error, contamination, and sabotage.

Immediate Cause. The immediate cause was hydrolysis of MIC due to water contamination. The exact source of the water has not been determined.

Contributing Causes:

1. Flare tower was shut down for repair.
2. Scrubber was inadequate to handle a large release.
3. Chilling system was turned off. (It also was too small.)
4. MIC tank was not equipped with adequate instrumentation.
5. Operating personnel lacked knowledge and training.
6. The inventory of MIC was large.
7. There was a lack of automatic devices and warning systems; it has been reported that safety systems had to be turned on manually.
8. When the plant was built, over 20 years before the accident, there were very few people near it. At the time of the accident, a shanty town had grown up near the plant with a density of 100 people per acre, greatly increasing the potential exposure of people to toxic releases. There was no emergency action plan to notify neighbors of the potential for toxic releases or of what to do if there was a release, nor was there a functioning alarm system.

Root Cause. The root cause of the accident appears to be a management system that did not adequately respond to the potential hazards of MIC. There was probably a greater inventory of MIC than was needed. The main process expertise was in the United States. Local management does not appear to have understood the process or the consequences of changes made. This includes plant design, maintenance and operations, backup systems, and community responsibility. The legal problems of Union Carbide are likely to continue for a long time. There appears to be enough blame to go around for all those involved in any way in the plant, including government units.

This accident has become widely known. It is an objective of many chemical process safety programs and government actions to "avoid another Bhopal"—that is, to avoid a severe release of toxic chemicals (usually referring to toxic chemicals in the air). Almost every chemical company in the world has been affected by this incident in one way or another, in the design and operation of chemical plants, in community action programs, and in the activities of such organizations as the American Institute of Chemical Engineers, the Chemical

Manufacturers Association, and many governmental units.

Phillips Explosion, 1989[57]

On October 23, 1989, at approximately 1300, an explosion and fire ripped through the Phillips 66 Company's Houston Chemical Complex in Pasadena, Texas. At the site, 23 workers were killed, and more than 130 were injured. Property damage was nearly $750 million. Business interruption cost is not available but is probably a very large figure.

The release occurred during maintenance operations on a polyethylene reactor. Two of the six workers on the maintenance crews in the immediate vicinity of the reactor leg where the release occurred were killed, together with 21 other employees of the facility. Debris from the plant was found 6 miles from the explosion site. Structural steel beams were twisted like pretzels by the extreme heat. Two polyethylene production plants covering an area of 16 acres were completely destroyed.

The Phillips complex produces high-density polyethylene, which is used to make milk bottles and other containers. Prior to the accident, the facility produced approximately 1.5 billion lb of the material per year. It employed 905 company employees and approximately 600 daily contract employees. The contract employees were engaged primarily in regular maintenance activities and new plant construction.

The accident resulted from a release of extremely flammable process gases that occurred during regular maintenance operations on one of the plant's polyethylene reactors. It is estimated that within 90 to 120 seconds, more than 85,000 pounds of flammable gases were released through an open valve. A huge unconfined flammable vapor cloud was formed that came into contact with an ignition source and exploded with the energy of 4800 lb of TNT. The initial explosion was equivalent to an earthquake with a magnitude of 3.5 on the Richter scale. A second explosion occurred 10 to 15 minutes later when two isobutane tanks exploded. Each explosion damaged other units, creating a chain reaction of explosions. One witness reported hearing ten separate explosions over a 2-hour period.

In the process used by Phillips at this site to produce high-density polyethylene, ethylene gas is dissolved in isobutane and, with various other chemicals added, is reacted in long pipes under elevated pressure and temperature. The dissolved ethylene reacts with itself to form polyethylene particles that gradually come to rest in settling legs, where they are eventually removed through valves at the bottom. At the top of these legs, there is a single ball valve (DEMCO® brand) where the legs join with other reactor pipes. The DEMCO valve is kept open during production so that the polyethylene particles can settle into the leg. A typical piping settling leg arrangement is shown in Fig. 5.14.

In the Phillips reactor, the plastic material frequently clogged the settling legs. When this happened, the DEMCO valve for the blocked leg was closed, the leg disassembled, and the block removed. During this particular maintenance process was going on, the reaction

Fig. 5.14. Typical piping settling leg arrangement.

block removed. During this particular maintenance process, the reactor settling leg was disassembled and the block of polymer removed. While this maintenance process was going on, the reaction continued, and the product settled in the legs that remained in place. If the DEMCO valve were to open during a cleaning-out operation, there would be nothing to prevent the escape of the gas to the atmosphere.

After the explosion it was found that the DEMCO valve was open at the time of the release. The leg to be cleaned had been prepared by a Phillips operator. The air hoses that operated the DEMCO valve were improperly connected in a reversed position such that a closed DEMCO valve would be opened when the actuator was in the closed position. In addition, the following unsafe conditions existed:

1. The DEMCO valve did not have its lockout device in place.
2. The hoses supplied to the valve actuator mechanism could be connected at any time even though the Phillips operating procedure stipulated that the hoses should never be connected during maintenance.
3. The air hoses connecting the open and closed sides of the valve were identical, thus allowing the hoses to be cross-connected and permitting the valve to be opened when the operator intended to close it.
4. The air supply valves for the actuator mechanism air hoses were in the open position so that air would flow and cause the actuator to rotate the DEMCO valve when the hoses were connected.
5. The DEMCO valve was capable of being physically locked in the open position as well as in the closed position. The valve lockout system was inadequate to prevent someone from inadvertently opening the DEMCO valve during a maintenance procedure.

Established Phillips corporate safety procedures and standard industry practice require backup protection in the form of a double valve or blind flange insert whenever a process or chemical line in hydrocarbon service is opened. According to OSHA, Phillips had implemented a special procedure for this maintenance operation that did not incorporate the required backup. Consequently, none was used on October 23.

The consequences of the accident were exacerbated by the lack of a water system dedicated to fire fighting, and by deficiencies in the shared system. When the process water system was extensively damaged by the explosion, the plant's water supply for fighting fires was also disrupted. The water pressure was inadequate for fire fighting. The force of the explosion ruptured waterlines and adjacent vessels containing flammable and combustible materials. The ruptured water lines could not be isolated to restore water pressure because the valves to do so were engulfed in flames. Of the three backup diesel pumps, one had been taken out of service and was unavailable, and another soon ran out of fuel. It was necessary to lay hoses to remote sites—settling ponds, a cooling tower, a water treatment plant, and so on. Electric cables supplying power to regular service pumps were damaged by fire, and those pumps were rendered inoperable. Even so, the fire was brought under control within 10 hours.

In the months preceding the explosion, according to testimony, there had been several small fires, and the alarm had sounded as many as four or five times a day. There had been a fatality at the same plant doing a similar operation about three months before this incident. Some of the employees in the area where the release occurred may not have heard the siren because of the ambient noise level, and may not have known of the impending disaster. Employees in the immediate area of the release began running as soon as they realized the gas was escaping.

The large number of fatalities was due in part to the inadequate separation between buildings in the complex. The site layout and the proximity of normally high-occupancy structures, such as the control and finishing building, to large-capacity reactors and hydrocarbon vessels contributed to the severity of the event.

The distances between process equipment were in violation of accepted engineering practices and did not allow personnel to leave the polyethylene plants safely during the initial vapor release; nor was there sufficient separation between reactors and the control room to carry out emergency shutdown procedures. The control room, in fact, was destroyed by the initial explosion. Of the 22 victims' bodies recovered at the scene, all were located within 250 feet of the vapor release point.

OSHA's investigation revealed that a number of company audits had identified unsafe conditions but largely had been ignored. Thus, a citation for willful violations of the OSHA "general duty" clause was issued to Phillips with proposed penalties of $5,660,000. In addition, proposed penalties of $6,200 were issued for other serious violations. A citation for willful violations with proposed penalties of $724,000 was issued to Fish Engineering and Construction, a Phillips maintenance contractor. Other financial penalties have been proposed. In the investigation it became apparent that Fish had become accustomed to tolerating safety and health violations at the site by its personnel and Phillips personnel, as well as participating in those violations by knowing about them and not taking direct positive action to protect its employees.

Since 1972, OSHA has conducted 92 inspections in the Dallas region at various Phillips locations; 24 were in response to a fatality or a serious accident. OSHA determined that Phillips had not acted upon reports by its own safety personnel and outside consultants who had pointed out unsafe conditions. OSHA also had conducted 44 inspections of the Fish Company, seven of them in response to a fatality or a serious accident.

One of the major findings by OSHA was that Phillips had not conducted a process hazard analysis or equivalent (such as HAZOP) in its polyethylene plants.

Immediate Cause. There was a release of flammable process gases during the unplugging of Number 4 Reactor Leg on Reactor 6 while undergoing a regular maintenance procedure by contractor employees. The unconfined flammable vapor cloud was ignited and exploded with devastating results.

The immediate cause of the leak was that a process valve was opened by mistake while the line was open. The valve was open to the atmosphere without a second line of defense such as another valve or a blind flange.

Contributing Causes to the Accident and the Severity of the Accident:

1. Procedures to require backup protection in the form of a double valve or a blind flange insert were not used. The lockout system was inadequate.
2. Air hoses were improperly connected in the reversed position.
3. The air hoses for the open and closed side of the valve were identical, allowing the hoses to be cross-connected.
4. The DEMCO valve actuator mechanism did not have its lockout device in place.
5. There was not a water system dedicated to fire fighting, and there were deficiencies in the system shared with the process.
6. The site layout and proximity of high-occupancy structures contributed to the severity.
7. There was inadequate separation of buildings within the complex. Especially, there was inadequate spacing between the reactors and the control room.

Root Causes. The root causes of the accident and its extreme severity appear to be failures of the management system, as shown by the following:[57]

1. According to OSHA, Phillips had not conducted a process hazard analysis or equivalent (such as HAZOP) in its polyethylene plants.
2. It was found by OSHA that the contractor, Fish Engineering, had a history of serious and willful violations of safety standard, which Phillips had not acted upon. The same contractor also had been

involved in a fatal accident at the same facility three months earlier.

3. A report by OSHA stated that Phillips had not acted upon reports issued previously by the company's own safety personnel and outside consultants. Phillips had numerous citations from OSHA since July 1972. OSHA discovered internal Phillips documents that called for corrective action but which were largely ignored.
4. Safe operating procedures were not required for opening lines in hazardous service.
5. An effective safety permit system was not enforced with Phillips or contractor employees, especially line opening and hot work permits.
6. Buildings containing personnel were not separated from process units in accordance with accepted engineering principles, or designed with enough resistance to fire and explosion.
7. The fire protection system was not maintained in a state of readiness:
   (a) One of the three diesel-powered water pumps had been taken out of service.
   (b) Another of the three diesel-powered water pumps was not fully fueled, and it ran out of fuel during the fire fighting.
   (c) Electric cables supplying power to regular service fire pumps were not located underground and were exposed to blast and fire damage.

SUMMARY

As the tragic case histories unfold, the significance of the process safety consideration presented in this chapter becomes chillingly apparent, and the necessity for inherently safe process design is revealed. The case histories also reveal significant flaws in management systems, which tends to be true for most chemical process accidents. Even an ideally safe process can be transformed into one with a high potential for disaster if a valid system is not in place to ensure that the inherently safe process design retains its integrity.

Because there is always risk when equipment, instrumentation, and human activity are involved, there is no method of making a plant completely safe. However, facilities can be made "inherently safer" by careful examination of all aspects of design and management, using modern techniques that are now available. If we are to improve our process safety performance and our public image, "inherently safe" process design coupled with "inherently safe" process management is imperative.

In addition to the information presented in this chapter and in the publications it has cited, references 58 through 72 are recommended as appropriate source material.

REFERENCES

1. M & M Protection Consultants, *One Hundred Largest Losses: A Thirty Year Review of Property Damage in the Hydrocarbon-Chemical Industry* 12th ed. 222 S. Riverside Plaza, Chicago, IL 60606, 1989.
2. Lees, F. P., *Loss Prevention in the Process Industries*, London, Butterworths, 1980.
3. Englund, S. M., "Opportunities in the Design of Inherently Safer Chemical Plants," *Advances in Chemical Engineering*, James Wei et al. (Eds.), Academic Press, New York, 1990.
4. Kletz, T., *Cheaper, Safer Plants or Wealth and Safety at Work*, Institution of Chemical Engineers, Rugby, England, 1985.
5. Schaller, L. C., Du Pont, Wilmington, DE 19898, *Plant Operations Progress*, No. 1 (Jan. 1990).
6. Howard, W., Consultant, Monsanto Chemical Company, Personal Communication, 1981.
7. *NFPA 68*, National Fire Protection Association, Batterymarch Park, Quincy, MA, 1988.
8. Bartknecht, W., *Explosions Course Prevention Protection*, Springer-Verlag, Berlin, 1981.
9. *NFPA 69* National Fire Protection Association, Batterymarch Park, Quincy, MA, 1986.
10. Liening, G., "Prolonging Service Life of Metals and Alloys in Chemical Processing Plants," *Chemical Processing*, p. 22 (Sept. 1986).

11. Liening, G., The Dow Chemical Co., Midland, MI, personal communication, (Sept. 5, 1986).
12. *Perry's Chemical Engineers' Handbook*, 6th ed., pp. 23–48, McGraw-Hill, New York, 1984.
13. Grinwis, D., Process Systems Associate; Larsen, Paul, Associate Instrument Engineering Consultant; and Schrock, Luther, Process Consultant, The Dow Chemical Co., Midland, MI and Freeport, TX, personal-communication, (Oct. 1986).
14. Fisher, H., *Chemical Engineering Progress*, p. 3 (Aug. 1985).
15. Huff, J. E., "A General Approach to the Sizing of Emergency Pressure Relief Systems," *Reprints of Int. Symp. on Loss Prev. and Safety Promotion in the Process Ind.*, Heidelberg, Germany, 1977, p. IV 223, DECHEMA, Frankfurt (1977).
16. Huff, J. E., *Institute of Chemical Engineers Symposium Series, No. 85*, p. 109 (1984).
17. Huff, J. E., "Emergency Venting Requirements for Gassy Reactions from Closed System Tests," *Plant/Operations Progress*, **3** (1), 50–59 (Jan. 1984).
18. Huff, J. E., "The Role of Pressure Relief in Reactive Chemical Safety," International Symposium on Preventing Major Chemical Accidents, Washington, DC, sponsored by the Center for Chemical Process Safety of the American Institute of Chemical Engineers, The United States Environmental Protection Agency, and the World Bank, (Feb. 3, 1987).
19. Fauske, Hans K., and Leung, J., "New Experimental Technique for Characterizing Runaway Chemical Reactions," *Chemical Engineering Progress*, p. 39 (Aug. 1985).
20. Jackson, B. L., Piping Specialist, The Dow Chemical Co., Midland, MI, personal communication, (Sept. 17, 1986).
21. Alexander, S., and King, R., Materials Engineering Dept., Michigan Div., The Dow Chemical Co., Midland MI, personal communication, (Feb. 14, 1988).
22. Cromie, J., Engineering Associate, The Dow Chemical Co., Midland, Michigan, personal communication, 1986.
23. Reynolds, J. A., Union Carbide Corp., "Canned Motor and Magnetic Drive Pumps," *Chemical Processing*, pp. 71–75 (Nov. 1989).
24. *Technical Management of Chemical Process Safety*, Center for Chemical Process Safety, American Institute of Chemical Engineers, 345 E. 47th St., New York, NY, 1989.
25. Cloud, M. J., "Fire, the Most Tolerable Third party," *Michigan Natural Resources*, p. 18, (May–June, 1990).
26. *Fire Safety Data*, Fire Protection Association, 140 Aldersgate St., London EC1A 4HX, 1988.
27. Bodurtha, F. T., Engineering Dept., Du Pont, *Industrial Explosion Prevention and Protection*, McGraw-Hill, New York, 1980.
28. Zabetakis, M. G., "Flammability Characteristics of Combustible Gases and Vapors," Bulletin 627, U.S. Dept. of the Interior, Bureau of Mines, Washington, DC (1965).
29. Cawse, J. N., Pesetsky, B., and Vyn, W. T., "The Liquid Phase Decomposition of Ethylene Oxide," Union Carbide Corporation, Technical Center, South Charleston, WV 25303 (no date available).
30. Stull, D. R., *Fundamentals of Fire and Explosion*, p. 50, Corporate Safety and Loss Dept., The Dow Chemical Co., Midland, MI, American Institute of Chemical Engineers, New York, 1976.
31. Kohlbrand, "Case History of a Deflagration Involving an Organic Solvent/Oxygen System Below Its Flash Point," 24th Annual Loss Prevention Symposium, Sponsored by the American Institute of Chemical Engineers, San Diego, CA, (Aug. 19–22, 1990).
32. Mackenzie, J., "Hydrogen Peroxide without Accidents," *Chemical Engineering*, pp. 84ff (June 1990).
33. Eichel, F. G., "Electrostatics," *Chemical Engineering*, pp. 154–167 (Mar. 13, 1967).
34. Boundy, R. H., and Boyer, R. F., *Styrene*, American Chemical Society Monograph Series, p. 63, Reinhold Publishing Co., New York, 1952.
35. *NFPA 77*, National Fire Protection Association, Batterymarch Park, Quincy, MA, 1986.
36. Bodurtha, F. T., "Industrial Explosion Control Course," Center for Professional Advancement, Chicago, IL, (Sept. 14–16, 1987).
37. Webley, P., Director, Massachusetts Institute of Technology Practice School, Midland Station, Midland, MI, personal communication (May 1990).
38. *CPQRA (Chemical Process Quantitative Risk Analysis)*, Center for Chemical Process Safety of the American Institute of Chemical Engineers, 1989.
39. TNO, "Methods for the Calculation of the Physical Effects of the Escape of Dangerous Materials: Liquids and Gases," ("The Yellow Book"), Apeldoorn, The Netherlands, 1979.
40. API RP 521 (2nd ed.), American Petroleum Institute, Washington, DC (1982).
41. Walters, S., "The Beginnings," *Mechanical Engineering*, **4**, 38–46 (1984).
42. Prugh, R. W., "Quantitative Evaluation of BLEVE Hazards," AICHE Loss Prevention Symposium, Paper No. 74e, AICHE Spring National Meeting, New Orleans, LA, (Mar. 6–10, 1988).
43. Moore, C. V., "The Design of Barricades for Hazardous Pressure Systems," *Nuclear Engineering Design*, **5**, 1550–1566 (1967).
44. Clancey, V. J., "Diagnostic Features of Explosion Damage," Sixth Int. Meeting of Forensic Sciences, Edinburgh, 1972.

45. Baker, W. E., Cox, P. A., Westine, P. S., Kulesz, J. J., and Strehlow, R. A., *Explosion Hazards and Evaluation*, Elsevier, New York, 1983.
46. Ludwig, E. E., *Applied Process Design for Chemical and Petrochemical Plants*, Vol. 1 (2nd ed.), Gulf Publishing, Houston, TX, 1977.
47. Brasie, W. C., Michigan Division, Process Engineering, The Dow Chemical Co., Midland, MI, personal communication, (Mar. 9, 1983).
48. Brasie, W. C., Michigan Division, Process Engineering, The Dow Chemical Co., Midland, MI, personal communication, (Oct. 6, 1982).
49. Allen, W. T., Process Engineering, The Dow Chemical Co., Midland, MI, personal communication, (May 1988).
50. AIChE, *Guidelines for Safe Storage and Handling of High Toxic Hazard Materials*, Center for Chemical Process Safety, American Institute of Chemical Engineers, 345 E. 47th St., New York, NY, 1988.
51. Knowlton, R. E., *Hazard and Operability Studies*, Chemetics International Co., Ltd., Vancouver, BC, Canada V6J, 1C7, (Feb. 1989).
52. NUS Corp., HAZOP Study Team Training Manual, *Predictive Hazard Identification Techniques for Dow Corning Facilities*, Gaithersburg, MD, (July 1989).
53. Technica Consulting Scientists and Engineers, London, England, 1990.
54. Technica, Inc., "HAZOP Leaders Course," 1989, Columbus, OH, Nov. 6–10; course leaders David Slater and Frederick Dyke.
55. Arendt, J. S., Lorenzo, A. F., and Lorenzo, D. K., "Evaluating Process Safety in the Chemical Industry," *A Manager's Guide to Quantitative Risk Assessment*, Chemical Manufacturers Association, DC. 1989.
56. Burk, Art., Principal Safety Consultant, Du Pont, Newark, DE, personal communication, (Feb. 20, 1990).
57. OSHA (Occupational Safety and Health Administration), U.S. Department of Labor, *The Phillips Company Houston Chemical Complex Explosion and Fire*, (Apr. 1990).
58. Bartknecht, W., *Dust Explosions Course, Prevention, Protection*, Springer-Verlag, Berlin, 1989.
59. Beveridge, H. J. R., and Jones, C. G., "Shock Effects of a Bursting-Disk in a Relief Manifold," Institution of Chemical Engineers Symposium, Series No. 85.
60. Burgess, D., and Zabetakis, M. G., "Fire and Explosion Hazards Associated with Liquefied Natural Gas," U.S. Bureau of Mines, Report RI 6099 (1962).
61. Burk, A., Principal Safety Consultant, Du Pont, Newark, DE, presentation on "Process Hazards Analysis and Quantitative Risk Assessment," (July 20, 1989).
62. "CEFIC Views on the Quantitative Assessment of Risks From Installations in the Chemical Industry," European Council of Chemical Manufacturers' Federations, Bruxelles, (Apr. 1986).
63. *Condensed Chemical Dictionary*, Van Nostrand Reinhold Co., New York, 1983.
64. *Guidelines for Vapor Release Mitigation*, Center for Process Safety, American Institute of Chemical Engineers, 345 E. 47th St., New York, NY, (Sept. 1987).
65. Hymes, I., "The Physiological and Pathological Effects of Thermal Radiation," UKAEA Safety and Reliability Directorate, Report SRD R275, Culcheth, UK.
66. Klein, H. H., "Analysis of DIERs Venting Tests: Validation of a Tool for Sizing Emergency Relief Systems for Runaway Chemical Reactions," *Plant Operations Progress*, **5** (1), 1–10, (Jan. 1986).
67. Nazario, F. N., Exxon Research and Engineering Co., *Chemical Engineering*, pp. 102–109 (Aug. 15, 1988).
68. Pieterson, C. M. and Huaerta, S. C., "Analysis of the LPG Incident in San Juan Ixhatepec, Mexico City, 19 Nov. 1984," TNO Report B4-0222, P.O. Box 342 7300 AH, Apeldoorn, The Netherlands (1985).
69. Roberts, A. F., "Thermal Radiation Hazards from Releases of LPG from Pressurized Storage," *Fire Safety Journal*, **4** (3), 197 (1981).
70. Townsend, D. I., and Tou, J. C., "Thermal Hazard Evaluation by an Accelerating Rate Calorimeter," *Thermochimica Acta*, **37**, 30 (1980).
71. Welty, J. R., *Engineering Heat Transfer*, SI Version, John Wiley & Sons, New York, 1978.
72. Wensley, J. H., "Improved Alarm Management through Use of Fault Tolerant Digital Systems," Instrument Society of America, International Conference and Exhibit, Houston, TX, (Oct. 13–16, 1986).

6

Managing an Emergency Preparedness Program

Thaddeus H. Spencer*

INTRODUCTION

Prevention, Prediction, and Preparation

The preceding chapter explored many technical aspects of chemical process safety and some safety management systems that form the foundation of a comprehensive emergency preparedness program. Clearly, the first step in preparing for emergencies is to identify and mitigate the conditions that might cause them. This process starts early in the design phase of a chemical facility, and continues throughout its life. The objective is to *prevent* emergencies by eliminating hazards wherever possible.

Although hazard elimination is the goal, experience has taught us that guaranteed, failure-free designs and devices have so far eluded human kind, despite astonishing advances in knowledge and technology. Even the most "inherently safe" chemical facility must prepare to control potentially hazardous events that are caused by human or mechanical failure, or by natural forces such as storms or earthquakes.

The process of careful, structured analysis and evaluation used to eliminate hazards during design and construction will also allow a chemical facility to accurately *predict* unplanned events that may create emergencies, and to effectively *prepare* to manage them should they occur. A comprehensive emergency preparedness program has all of these elements: *prevention*, *prediction*, and *preparation*.

The fundamental need to predict and prepare for a failure of some kind is familiar to everyone. Fortunately, most of the failures that we encounter create little more than inconveniences in our lives. Others have much more serious potential. Such a failure can trigger an emergency, a term that Webster defines as "an unforeseen combination of circumstances or the resulting state that calls for immediate action." If the immediate action is ineffective, the emergency will escalate to a full-blown crisis.

Certainly most if not all of us in our personal experiences have had many opportunities to reconfirm the wisdom of the admonition,

*Senior Consultant, Safety and Environmental Management Services, E. I. du Pont de Nemours and Company.

"Plan for the best, but prepare for the worst." As a result, we prepare ourselves for human and mechanical failure in a variety of ways, some so simple and familiar that we scarcely are aware that we are managing a personal emergency preparedness program.

To varying degrees, each of us has assessed our personal vulnerability to a specific emergency and the potential consequences to family and property. In some cases the required immediate action may be minor, but elsewhere the same combination of circumstances will demand significant resources to avoid a crisis of major proportions. The loss of household electric power in an urban condominium may only turn out the light, but on a farm, where electricity pumps the well water, milks the cows, refrigerates the produce, and irrigates the land, loss of power can bring disaster to the unprepared.

In an industrial environment, the consequences of human or mechanical failure can be far greater, even threatening the lives of employees and neighbors. Therefore, emergency prevention and preparedness efforts must have high priority, receiving continuous attention from every employee, including those in every level of management and supervision.

In the chemical industry, emergency preparedness programs have long been recognized as vital elements in protecting people, property, and the environment from harm. Few chemical facilities are without an emergency response plan. Still, when the alarms sound and the emergency is real, the response often does not proceed as planned. Too often, preparation for the unlikely event has been inadequate. As a result, most chemical facility managers are taking a fresh and critical look at their existing emergency preparedness programs in cooperation with their communities, and placing a still higher priority on being truly prepared for emergencies. This is consistent with the aggressive approach to other aspects of safety management that has made the chemical industry one of the safest industries.

Although the commitment to *prevent* the events that lead to injuries and emergencies of all kinds remains the first priority in safety management, we must "prepare for the worst." In this chapter we will explore how chemical facilities and their neighbors can better *predict* and *prepare* for unplanned events that threaten lives, property, and the environment—events that call for immediate and coordinated action.

Need for Emergency Preparedness Programs

The need for more effective management of emergency preparedness programs by chemical facilities and their host communities became painfully clear to the world in the 1980s—a decade marred by tragic events linked to the manufacture and distribution of chemical products. These events revealed serious deficiencies, not only in training people to react effectively during an emergency, but also in managing the systems employed to identify, evaluate, and mitigate hazards that may cause an emergency.

The chemical industry has vigorously responded to these problems in a variety of ways. In some cases well before the crisis of the 1980s, individual chemical units took aggressive action to increase the reliability of their operations, to communicate these actions to their neighbors, and to involve the communities in the process of emergency response planning. This proactive interaction at the local level proved fruitful in promoting the active partnership needed to solve mutual problems and to respond to public concerns. In 1986, these individual initiatives were institutionalized by the industry in the Chemical Manufacturers Association's "Community Awareness/Emergency Response (CAER)" program, which has in turn become a major element of the more recent "Responsible Care" program.

The proven success of these activities provides a powerful reason for giving high priority to open and cooperative management of an emergency preparedness program. By implementing such a program, a chemical facility builds important bridges to its neighbors and fosters positive perceptions. Involving

the host community provides an opportunity to demonstrate a sincere commitment to the protection of people and the environment, and a route to the mutually profitable solution of real problems through meaningful two-way communication with the public.

Much of what had been voluntary prior to 1986 became law in the United States with the enactment of the Superfund Amendments and Reauthorization Act (SARA). The Emergency Planning and Community Right-to-Know portions of the Act (known as Title III, or EPCRA) place specific organizational, planning, communication, and training responsibilities on the public and private sectors, as do the accidental release provisions of the Clean Air Act amendments enacted in 1990. Other laws also require emergency response plans. In addition various state and local regulations must be considered, understood, and reflected in a complete emergency preparedness program.

These laws and regulations, which will be covered later in this chapter, should be viewed as describing only the minimum requirements. Beyond the legal and self-protective reasons for chemical facilities to improve their management of emergency response programs lie other important driving forces. One of these is the moral and ethical responsibility to employees and the public to work toward elimination of events destructive to the quality of life. In practice, full recognition of this responsibility requires actions in prevention, prediction, and preparation that go beyond the letter of the law.

The financial benefits of good emergency preparedness program management are important as well. Emergencies are always expensive. Uncontrolled emergencies can become financially devastating crisis. Nervous communities, fearful of the dangers dimly perceived through industrial fences, can and have put some chemical operations out of business. Creating and maintaining a comprehensive emergency preparedness program does not come free. However, the investment is proving to be money well spent by most chemical facilities.

PREVENTING AND PREDICTING EMERGENCIES: GETTING STARTED

The objective of a comprehensive emergency preparedness program is the protection of people, property, and the environment from unplanned hazardous events. Organizations handling hazardous materials recognize that the process of creating an effective program starts with the identification, evaluation, reduction, and control of hazards (commonly called risk management), and proceeds through the preparation, drilling, and maintenance of plans and procedures designed to contain an emergency situation should one occur.

To assure an effective emergency preparedness program, chemical facilities need to make periodic, formal assessments of their vulnerability to and preparedness for emergencies. Managers must be involved in developing and monitoring key indicators that will help in assessing an organization's ability to prevent or deal with an emergency. Awareness of any program weaknesses revealed by examining these indicators leads to corrective action to ensure that the potential for incidents decreases, and emergency preparedness improves. Sample checklists including some of these key indicators are shown in Fig. 6.1.

Central to all emergency preparedness programs is a written emergency response plan (ERP). An ERP obviously is a key element of emergency preparedness, yet it is only one procedural part of a systematic process that includes the following sequential steps:

- Identify and evaluate hazards.
- Mitigate hazards wherever possible.
- Identify and evaluate remaining hazards.
- Identify and evaluate resources.
- Develop emergency procedures and ERP.
- Train facility personnel.
- Communicate plans to public.
- Integrate with community ERPs.
- Conduct and critique drills.
- Review and revise ERPs.
- Do it again.

EMERGENCY PREVENTION	EMERGENCY RESPONSE PLANNING
1. Frequent management presence in operating areas.	1. Emergency manual that documents the areas or processes with emergency potential and describes the emergency response plan.
2. Proper storage and identification of hazardous materials.	2. Regular emergency response plan training, and drills each quarter.
3. Formal and systematic inspections of key equipment, safety devices, and safety interlocks.	3. Liaison with the community for response coordination.
4. Audits for compliance with safety rules and procedures.	4. Emergency response management organization with documented functional tasks and assigned personnel.
5. Periodic critical reviews of existing operating and maintenance procedures.	5. Systems to recognize and report an emergency in a timely manner.
6. Training programs updated to meet current needs.	6. Documentation of hazardous chemicals with potential to leave the site, and evacuation plan for affected areas.
7. Control systems for maintenance hot work and changes in process and equipment.	7. Adequate emergency response equipment.
8. Systematic process hazards reviews during design, start-up, and routine operations.	8. Proper method to account for personnel during and after an emergency.
9. Formal reviews all of events that could or did cause serious process incidents.	9. Procedures to review and modify the emergency plan following drills or actual emergencies.

Fig. 6.1. Emergency preparedness checklists.

The creation and maintenance of an ERP requires the allocation of valuable resources, as well as strong leadership from a manager. No one else in the organization has the authority to commit the resources required. Only the power of the manager's office can overcome the organizational inertia that is often encountered.

Many people do not like to "prepare for the worst." They may harbor sincere doubts about the value of planning for events that they consider unlikely to occur. Some people are complacent after years free of serious accidents, and honestly believe that "It can't happen here." Others may be concerned that an open discussion of potentially disastrous events will needlessly upset employees and neighbors. These and other "start-up" problems could make trouble for a manager initiating (or reviving) an emergency preparedness program.

Managers need to find ways to stimulate employee interest and enthusiasm in the planning process. This is best done through the involvement of employees. One strategy that minimizes problems is the early involvement of those employees who serve as emergency responders in their communities.

Some of them have witnessed the tragic consequences of failing to heed early danger signs and being unprepared for events. Thus, they can be willing and able catalysts in the emergency response planning process.

Open communication with all employees early in the process also is important. Sharing objectives and encouraging contributions stimulates thoughtful discussion, leads to more complete hazard recognition and mitigation, and ensures a greater chance of effective action when the emergency alarms do sound. Armed with facts, employees acting as informal ambassadors to the community can reduce the potential for public alarm over a facility's preparations to deal with serious but low-probability events.

HAZARD IDENTIFICATION AND MITIGATION

As shown in Fig. 6.2, emergency preparedness begins with the identification and mitigation of hazards. When properly done, the systematic analysis and evaluation of chemical process hazards stimulate actions that eliminate the potential for many emergency situations and pinpoint the situations that remain. The direction of these analytical and corrective efforts is an integral part of managing

Fig. 6.2. Elements of a comprehensive emergency preparedness program.

a comprehensive emergency preparedness program.

Process Safety Management Team

Hazard identification and mitigation are the responsibility of a facility's line organization—the same people who are responsible for all other organizational performance parameters. However, a line organization often needs support in identifying and mitigating complex chemical process hazards, support that can be effectively provided through the formation of a standing team functioning under the manager's direction. This team, called a Process Safety Management Team (PSMT), includes representatives of each unit of a facility from various levels of the organization. They meet on a regular basis (usually monthly) and report frequently to the manager on the status of their activities.

Identifying Hazards: PSR Teams

Supported by the PSMT, the line organization develops a structured approach for performing process hazard analyses on a repetitive basis. One effective way to do this is to organize knowledgeable facility personnel into ad hoc Process Safety Review (PSR) teams. These teams, supplemented by outside specialists as required, are responsible for studying all processes, identifying all potential hazards, and recommending appropriate corrective or control measures.

The makeup of a PSR team is critical to the success of this process. The effectiveness of the team depends on the skills, knowledge, and cooperative effort of its members, and the leadership ability of its chairperson. Each member must be familiar with the process being studied, and have at least a working knowledge of the basic engineering principles and chemistry involved. The team should include supervisors from operations and maintenance and a technical support person. Knowledgeable process operators and maintenance personnel would be valuable team members as well. Others who could contribute to the team's work include design engineers, specialists in electrical and instrument systems, safety engineers, reaction kinetics consultants, and equipment vendors.

The selection and training of PSR team members must be carefully monitored. Most managers make this the responsibility of the standing PSMT, which also may assist a PSR team in choosing the most appropriate hazards evaluation method for a specific process.

Review Methods

There are many structured methods for carrying out effective reviews of process hazards. The four most commonly used methods are:

- What if/checklist
- Failure mode and effect analysis
- Hazard and operability study (HAZOP)
- Fault tree analysis

What If/Checklist. The most frequently used method of process hazard review, the what if/checklist, is effective in reviews of relatively uncomplicated processes from raw materials to final product. The team formulates and answers "What if?" questions at each handling or processing step to evaluate the effects of component failures or procedural errors. They use a checklist to ensure that all important subjects are addressed. This method should be used as the first step in all process hazard reviews.

Failure Mode and Effect Analysis. When the team studies a specific item of equipment, such as a reaction vessel, they often use the failure mode and effect analysis method. Its semi-quantitative approach assists in prioritizing hazards.

Hazard and Operability Study (HAZOP). The HAZOP procedure systematically questions every part of a process to discover how deviations from the intention of the design can occur, and to determine if the consequences of such deviations are hazardous.

Fault Tree Analysis. Fault tree analysis, the most complex of the commonly used methods, is employed to determine the possible causes of a preselected undesired event. Through the use of logic diagrams and failure rate data, the team can make a quantitative evaluation of the frequency of the undesired event.

For additional information on such methods, refer to Chapter 5.

Recommendations and Reports

Regardless of the method used, the PSR team's most important responsibility is to alert management to serious hazards that may have been overlooked or given inadequate attention. To fulfill this responsibility, a PSR team must take the following steps:

1. Identify the hazards that could cause explosion, fire, release of toxic materials, serious injury, or inappropriate exposure to chemicals.
2. Evaluate the magnitude of the hazards for the areas of probable involvement; the consequences of an event in terms of injuries, environmental harm, and property damage; and, qualitatively or quantitatively, the probability of the hazards' occurrence.
3. Develop practical recommendations to eliminate or control the hazards identified.

The PSR team reviews in detail up-to-date reference material such as:

- Architectural drawings.
- Equipment layout drawings.
- Process schematics.
- Instrument diagrams.
- Chemical and physical characteristics of process materials.
- Equipment design specifications.
- Operating procedures.
- Process conditions.
- Emergency shutdown procedures.

When a facility has more than one chemical process, the PSMT develops a priority order for reviews and recommends a review frequency to the facility manager. The frequency usually ranges from once every two to three years for high-hazard-class processes such as explosives manufacture or acetylene purification to once every five to seven years for low-hazard-class processes such as alcohol purification, steam generation, and operations involving combustible materials.

Changes not anticipated in the original design of equipment often pose serious problems. Some examples of such changes are:

- Introduction of different raw materials.
- Changes in temperatures, pressures, speeds.
- Deterioration of equipment.

At the completion of each PSR, the team prepares a written report that defines needs, makes recommendations to remedy problems, and recommends priorities for the correction of deficiencies. Following review and acceptance of the report, the facility manager assigns responsibility for corrective action to the appropriate operating personnel. The PSMT then assists the manager in monitoring the status of the recommendations from all PSRs.

Mitigating Hazards: Release Detection and Mitigation

The release of flammable or toxic chemicals from uncontrolled pressure relief vents or as a result of equipment failure may present a serious threat to employees or neighbors who fail to guard against exposure. Every chemical facility must address this potential problem and prepare to protect people from these hazards by working toward reducing the potential hazard and its consequences. Typically, the line organization and PSR teams are responsible for hazard mitigation. They conduct a hazard study that includes the following activities:

- Detailed appraisal of the potential for accidental release of toxic gas or vapor.
- Evaluation of instruments and other methods for detecting such leaks.

- Provisions for rapid alerting of threatened personnel, and for communicating with emergency responders.
- Identification of buildings in which people might be trapped by such a release.
- Assessment of the capability of buildings or rooms to prevent the infiltration of gas or vapor.
- Evaluation of plans for building evacuations, including the provision and maintenance of appropriate personal protective equipment.

Early *detection* of a chemical leak is necessary in order to limit its effect on people and the environment. The most fundamental method for detecting a chemical release is the systematic patrol of operating units by personnel trained to recognize potentially hazardous vapors using odor or visual observations. The frequency and scope of the patrols will vary with the nature of the process equipment and materials; however, every facility should have documented patrol procedures in place.

When particularly hazardous vapors are present, the patrol procedures should include special provisions to protect the patrollers from the fumes. For example, in facilities producing or consuming hydrogen sulfide, operators may carry emergency respiratory equipment to escape any unexpected fumes that they may encounter, and patrol in pairs or individually, under constant surveillance.

Many chemical facilities supplement operator patrols with an instrumented detection system. Such a system may be a necessary resource when a hazards study concludes that the system will substantially increase the available escape or emergency response time, or where:

- The harmful substance is odorless or deadens the sense of smell at hazardous concentrations.
- The harmful substance is toxic at concentrations undetectable by smell.
- Large numbers of people may be exposed quickly.
- Ventilation systems might draw toxic fumes into a building before other means of detection could trigger protective action.

At the core of an instrumented leak detection system is a gas detector. There are many kinds of detectors on the market with varying degrees of sensitivity and selectivity. All require careful regular testing and maintenance. An engineering study of many site-specific factors is required before one makes a choice and designs an appropriately instrumented system. Some generally accepted guidelines are:

- Early warning of a leak is enhanced if the sensors can be placed near the process equipment subject to leakage.
- Air movement characteristics are critical in achieving reliable detection.
- Most detectors respond to several gases or fumes, so the possible presence of all airborne substances affecting the detection system must be considered.
- When it is necessary to monitor work areas or the intakes to ventilation systems, a highly sensitive system is desirable.

Most detection systems are designed not only to report the presence of hazardous fumes through instrument readouts but to sound an alarm and automatically initiate corrective or protective action. In an office or shop, for example, the system can be designed to shut down all ventilating fans and close exterior air inlet dampers.

When process safety reviews have identified chemical releases as potential sources of facility emergencies, the organization must provide the training and materials needed to ensure a prompt and appropriate reaction to *mitigate* the hazards. Some countermeasures that are effective in limiting the spread of a hazardous material spill or release should be included in the design of chemical process equipment, and described in emergency response procedures. The following discussion of release mitigation is largely excerpted from

Guidelines for Vapor Release Mitigation, prepared by R. W. Prugh and R. W. Johnson for the Center for Chemical Process Safety of the AIChE (copyright 1988 by the American Institute of Chemical Engineers, reproduced by permission of the Center for Chemical Process Safety of the AIChE).

"Water, steam, and air curtains and water sprays are primarily effective in dispersing and/or diluting vapors with air to reduce the severity of effects of a hazardous vapor release. In some cases, vapors can be partially 'knocked down' or absorbed after release."

"Ignition source control and deliberate ignition are also possible vapor release countermeasures." "For areas around processes handling flammable vapors, ignition source control is practiced to reduce the probability of vapor ignition if a leak occurs." "Administrative controls are exercised on plants where flammable materials are processed. Such controls may include hot work permits, restricted smoking areas, not allowing lighters or matches on the site, and electrical grounding and bonding procedures."

"Deliberate ignition is a countermeasure against spills of highly toxic materials which are also flammable, such as hydrogen sulfide, hydrogen cyanide, and methyl mercaptan. Igniting nontoxic flammable materials such as hydrocarbons may present hazards outweighing possible advantages." In any case, deliberate ignition must be carefully planned and executed so that the resulting fire is truly controllable.

"Practical methods for combating vapor from liquid leaks are dilution, neutralization, or covering. All three reduce the vaporization rate of the pool. Water dilution is effective for spills of water-miscible or water-soluble material. Spraying water into the spill reduces the vapor pressure by reducing the concentration of the liquid." "For acidic spills, limestone or soda ash is often used" to react with the spilled liquid to produce a less volatile salt or ester. "A foam cover can be effective in reducing vaporization from spills," and "dense liquids can be covered with lighter immiscible nonreactive liquids" to accomplish the same thing.

PREPARING FOR EMERGENCIES: IDENTIFYING AND EVALUATING RESOURCES

The process of hazard identification, evaluation, and mitigation, when sustained over a period of time and coupled with other good safety management practices, can prevent most process-related emergencies. However, the potential for various low-probability process-related events will remain. Should one occur, there must be resources available to promptly bring the event under control. These resources include designated personnel, plans, systems, and facilities that are needed for effective action and communication. The facility line organization is responsible for providing and maintaining these resources.

Personnel

Emergency Preparedness Team. As with the PSMT and the PSR team, the efforts of the line will benefit from the support of a standing emergency preparedness team functioning under the direction of the facility manager. This team, which includes representatives of each unit of the facility, leads the development and maintenance of an emergency response plan (ERP) for the facility, and monitors specific procedures and training for handling emergency situations of all kinds. It should meet on a regular basis (usually monthly), and report frequently to the manager on the status of its activities.

Emergency Management Organization. Emergencies demand rapid, well-coordinated decisions, communications, and action to bring them under control as swiftly as possible. There must be a formal emergency management organization (EMO) whose purpose is to achieve this objective. The structure of the organization is not critical as long as it is capable of rapid assembly, carried out its assigned responsibilities, and meets the needs of the facility. A generic EMO is shown in Fig. 6.3; the EMO is described below, in the section on "Developing an Emergency Response Plan."

Fig. 6.3. Generic emergency management organization.

Fire Brigades. Most chemical facilities of any size have established fire brigades as a key resource in their emergency preparedness programs. By virtue of its training and familiarity with the equipment and the physical layout of a site, the brigade can be a major factor in *preventing* incidents as well as in minimizing injuries and property damage due to fire, explosion, or other causes.

The size and the structure of a fire brigade vary, depending upon the hazards present, the size of the facility, and the level of internal fire-fighting capability desired; and large, multi-process facilities usually have several unit or area brigades. Each brigade has a designated captain and an alternate. The training of brigade members must be commensurate with the duties and the functions that they are expected to perform, and thus depends on the fire prevention plan option selected by site management. (Fire prevention plans are discussed later in this section.)

It is best to have a written statement establishing the brigade and its duties, specifying its size and organizational structure, and outlining the type, amount, and frequency of training provided. In the United States, OSHA has promulgated minimum standards for fire brigades, which may be found in 29 CFR 1910.156.

Plans

Work Unit Plans. In a large facility, each work unit has emergency procedures for its own area of operations, and these work unit plans form the foundation of the facility emergency response plan (ERP). Thus, the facility's emergency response can be no more effective than the recognition and response capabilities of the employees in the facility work units.

The primary objectives of these unit plans are (1) to control and contain any emergency condition within the unit, and (2) to provide protection for unit personnel and equipment from events originating outside the unit. Unit plans specify who has authority to take emergency action, and how escalation to a full facility emergency occurs. Subjects that must be covered include:

- Unit emergency command and coordination.
- Communications within the unit.
- Communications with other units.
- Emergency assessment.
- Unit evacuation and personnel accounting.
- Emergency shutdowns.
- Communications with the facility management.
- Criteria and procedures for securing resources from outside the unit.

Unit plans clearly assign primary responsibility for initial emergency assessment and reaction to the lead unit employee on site at the time of the event. Among the factors to be considered by that individual in making

the initial assessment are:

- Previous experience with similar situations.
- How long the situation has existed.
- What might occur "If...."
- Properties and hazards of materials involved.
- Other complicating situations.

Using the information gained from the assessment, unit personnel take whatever immediate action is required to protect people and property on and off the site.

Work unit plans should contain guidelines for assuring uniformity in the assessment of and reaction to unusual events. An effective approach is to develop an emergency classification system that includes criteria for classification and guidance for the appropriate response actions for each level of emergency that is defined.

For example, an emergency judged to be controllable within the unit, with no evacuation outside the unit and no impact in the community, is a Level One Emergency. Small releases of hazardous materials usually fall into this category. Appropriate response items to consider would include:

- Protection of unit personnel.
- Shutdown or isolation of affected equipment.
- Notification on- and off-site.
- Containment and cleanup.

Similarly, other events are categorized by their potential or actual severity. Preplanned response checklists are prepared for each one. An incident that requires response by facility personnel outside the affected unit but has no impact on the community is a Level Two Emergency. A Level Two Emergency requires activation of the facility ERP. If some community impact becomes probable, the emergency is a Level Three Emergency. That level of emergency triggers additional response requirements, including interaction with the community.

The unit plan must be formal and in writing. Because the emergency procedures of each individual operating area or work unit form the foundation of the facility ERP, they should be reviewed by the emergency preparedness team for completeness and consistency across the site before they are included in the facility ERP.

Fire Prevention Plan. One of the most basic resources that all chemical facilities must have is a fire prevention plan. To be fully effective, this plan must be in writing, and it must be reviewed with all employees on an established schedule.

A comprehensive fire prevention plan includes the following subjects:

- Major facility fire hazards.
- Storage and handling practices for combustible and flammable materials.
- Identification of potential ignition sources.
- Procedures to maintain systems and equipment installed to control ignition sources.
- Names and titles of those responsible for the system and equipment maintenance and the control of fuel source hazards.
- Procedures to minimize accumulations of flammable and combustible waste materials.
- Plans for communicating to all employees the fire hazards and their specific responsibilities in the event of fire.

Fire response options open to facility managers range from a plan to evacuate all employees without attempting fire fighting on any scale, to full involvement of an established fire brigade in fighting advanced-stage structural fires. The fire prevention plan must specify which option applies to the facility. The option selected determines the type and extent of education and training required, as well as the type of equipment needed on the site.

Fire prevention and response information may be covered in other written facility documents (such as job descriptions, rules, or procedures) and communicated to employees on the same schedule as the fire prevention plan.

Evacuation Plans. Many emergency events require the evacuation of at least some facility personnel, if only for precautionary reasons. This presents other needs, which must be addressed in the evacuation plan. Personnel who are directed to evacuate their normal work stations need visible wind direction indicators to help them move away from a dangerous gas cloud drifting downwind of the release point. If the building is enveloped in the cloud, however, they will need personal protective equipment in order to evacuate it safely. When there are processes that must be rapidly shut down and/or isolated as evacuation proceeds, the plan must include written procedures and appropriate means to ensure incident-free action.

Headcount Plan. The evacuation plan also must establish gathering points for all personnel in order to identify those who are missing. The procedure for accounting for personnel, called the headcount, must work effectively and rapidly in the first minutes of the emergency to determine if anyone is missing or known to be injured. Communications to the personnel chief or another designated member of the facility EMO must be quick and accurate so that search and rescue operations and medical aid can be successfully implemented.

Each work unit in a facility must have its own plan for headcounts as a part of its emergency preparedness program. This plan must be in writing, and it must be practiced frequently enough to ensure good execution in a time of real emergency. It should include:

- A designated assembly point for each employee.
- An alternate assembly point, should the primary location be inaccessible.
- A procedure for counting at each assembly point.
- A designated unit headcount coordinator and backup.
- Primary and alternate phone numbers where coordinator may be reached.
- Phone numbers for reaching the personnel chief, who coordinates and summarizes the site-wide accounting.

Headcount procedures must include guidelines for accounting for visitors to the facility at the time of the emergency. These visitors may include truck drivers, vendors, and contractors. Some guidelines are:

- All visitors must sign in and out, preferably at one designated gate.
- Visitors will have a designated host employee responsible for their safety whenever possible.
- Everyone temporarily on the facility must report to the assembly point of the unit he or she is in at the time of the emergency.
- Unit procedures must include reporting these "extra" people by name to headcount headquarters.
- The personnel chief of the EMO must have access to the log at the visitors' gate.

Alarm Systems

Another essential resource that every chemical facility must have is an effective alarm system, which initiates action by endangered personnel and emergency responders and serves the needs of people off as well as on the site.

A satisfactory facility emergency alarm system must meet four additional requirements. It must, at a minimum, immediately alert *all* the people on the site. (A single signaling device should be adequate for small sites; however, large facilities may require many devices placed to reach all occupied areas of the site.) The system must function even when the facility has a general power failure. Alarm activation controls should be located so that an emergency condition is unlikely to prevent access to them. Finally, an emergency warning system must be integrated with the surrounding community, and must be designed with the participation of local officials and special-needs groups.

On most older chemical facilities, the basic component of the emergency warning system is the fire alarm. It is essential that there be different signals to distinguish between an actual fire and other threatening events, as the action taken in response to a toxic gas release, for example, may be quite different

from that taken in response to a fire. On multi-process sites, the signals also should identify the work unit involved.

Depending upon the number of people at a facility and their familiarity with response and evaculation procedures, it may be necessary to supplement signals with verbal instructions delivered by public address equipment, radios, or automated telephone systems. For reliability at any time of day, a system using verbal instructions must be located in a regularly manned job station such as a control room or gate house.

Any alarm system used must communicate clearly the nature of the emergency event and its severity. In a Level One Emergency some people may have to be evacuated in order to ensure their safety, but by and large the emergency can be contained within the unit. In that case, information usually may be given to unaffected employees through normal telephone communications.

In the event of a Level Two or a Level Three Emergency, each facility must have an easily recognizable signal that clearly communicates that a major incident has occurred. Some facilities have alarm boxes that will automatically sound this signal when pulled. Others have boxes that require manual operation. Still others rely on a telephone message to trigger the signal.

In some facilities, the "major" emergency alarm is a steady blast on the facility steam whistle. Others use sirens or special bells. Whatever is chosen, the alarm must reach everyone on the site, triggering activation of the facility ERP and the initiation of a headcount procedure.

It is essential that every every employee on a chemical facility know how, where, and when to turn in an appropriate and effective emergency alarm. The specifics will vary considerably, but the procedure for turning in an alarm is similar in all cases.

Employees first need to know how to report a fire. On most facilities, this is done by using a fire alarm box or by telephoning a central station such as a guard house from which the fire alarm can be sounded. Chemical facilities need to have well-developed training and drill programs in place to be sure that response procedures and equipment are maintained.

When fire alarms are supplemented by verbal instructions, the facility personnel who are contacted need to know the location and the nature of the fire, the action under way to control it, and the actions required (if any) by those not involved. There are many commercially available communications systems that deliver this information efficiently, even on large sites.

Facilities for Protection and Communication

Safe Havens. When accidental releases of a toxic gas occur or threaten to occur, the immediate protection of on-site personnel is of paramount importance. One resource available for this purpose is a building or room that by the nature of its construction and its heating and ventilating characteristics can prevent the infiltration of intolerable concentrations of the toxic substance. The best location for a safe haven is determined by an engineering study. That location then is documented in the work unit and facility emergency response plans.

Safe havens may be rare on some sites; however, many buildings can provide personnel with temporary protection until the incident has been analyzed and a decision made on the need for evacuation. To qualify as a designated temporary safe haven, a building must be reasonably well sealed against air infiltration, with adjustable ventilation systems that can reduce or close off exhaust vents and outside makeup air. Emergency procedures should state how long a building can be considered a safe haven under specified exposure conditions.

Main Emergency Control Center. Using a list of safe havens developed by the work units the emergency preparedness team helps management select one building or room to be used as a control center in case of a major emergency. This Main Emergency Control Center (MECC) should be as remote from potential hazards as possible in order to serve

as a reasonably safe haven to those involved in directing, coordinating, and communicating activities for the duration of an emergency.

MECCs need to have basic information readily available and maintained in an up-to-date, standby condition. Such a collection would include:

- Copies of the facility ERP.
- Facility maps and diagrams.
- Process material isolation points.
- Fire control maps and diagrams.
- Maps of the surrounding area.
- Aerial photos of the facility and the surrounding area.
- Names, addresses, phone numbers for:
  - All facility employees.
  - Off-site company people to be notified.
  - Groups and organizations who may be notified.
  - Community officials who may be notified.
- MSDS (material safety data sheets) for facility materials.
- MSDSs for the materials of neighboring chemical facilities.
- Copies of mutual aid agreements.
- Highlights of the facility's history, products, and performance.

The MECC should have the following basic equipment:

- Adequate telephones and lines.
- Unlisted telephone and/or a hotline.
- Two-way radios.
- Fume path projector.
- FAX machine.
- Regular radio with tape recorder.
- Regular television with recorder and playback video cassette recorders.
- Chart pads and stands.
- Battery-powered lighting.
- Personal protective equipment.
- Hand-held tape recorders.

Alternate Control Centers. Recognizing the unpredictable nature of a developing emergency, it is advisable to select at least one alternate on-site main control center. Both the MECC and the alternate on-site center should have backup power supplies.

The presence of substantial quantities of explosive or toxic materials on-site may justify establishing yet a third location off-site, in case of a complete evacuation of the facility. This could be a fixed location in the nearby community, or a mobile unit such as a truck or van properly equipped for managing an emergency. Some of the supplies and equipment for the alternate off-site main control center may have to be packed in a readily transported kit rather than being on standby at the alternate off-site location.

Media Headquarters. In an emergency, it is essential to maintain regular contact with the public and the media. A separate location for communications will allow that contact to continue without interfering with the operations of the MECC or overloading its communications system. There, the public affairs chief and the facility manager can receive media representatives and provide periodic updates on the emergency situation. The media headquarters should be equipped with several direct outside telephones and the facilities to make the reporters comfortable for the duration of the emergency. Permanently mounted facility plot plans or aerial photographs are helpful resources, as are handouts describing in general terms the facility and its products.

DEVELOPING AN EMERGENCY RESPONSE PLAN

Following the identification and evaluation of the resources available to the facility, the emergency preparedness team, under the direction of the manager, develops the written emergency response plan (ERP) for the entire facility. The emergency preparedness team is responsible for ensuring that the written facility ERP informs all employees of their roles in an emergency and the hazards to which they may be exposed.

An ERP must be tailored to a specific facility, reflecting its unique conditions and individual needs. Among the variables that

affect the details of a plan are:

- Materials used, produced, or stored.
- Nature of the operations.
- Available employee skills.
- Geographic location.
- Proximity to other facilities.
- Available emergency resources.
- Mutual aid agreements.

No two ERPs will be exactly alike, nor is there one best outline. The plan of a production facility will differ from that of a warehouse. However, every plan must include:

- Initial alarm procedures.
- Emergency escape procedures and routes.
- Emergency operating and shutdown procedures for critical operations.
- Accounting procedures for all on-site personnel.
- rescue and medical duties for specified personnel.
- Procedures for communicating the emergency inside and outside the organization.
- Structure, duties, and resources of the EMO, including names or titles of people with detailed knowledge of the plan and its assigned duties.
- Reference material.

Plan Design

The ERP is a working plan, applicable to any event with emergency potential occurring at any time and at any location on the facility. It should be written concisely, with diagrams and checklists used wherever possible, so that it may be effectively used for guidance during an actual emergency. To accomplish this, it is helpful to divide the ERP into several major sections, which are in turn divided by subject.

For example, an ERP may be broken down into three sections: the Emergency Management Organization, Action Plans, and Reference Material.

The first section documents the structure, duties, resources, and communications systems for the facility Emergency Management Organization (EMO), and the conditions that will trigger the EMO's response. It includes plans for site-wide notification and response.

The second section includes detailed action plans for each particular type of emergency. This section includes summaries or outlines of the emergency procedures developed by each work unit of a facility, and may contain considerable process-oriented information.

The third section consists of reference material on plan philosophy, training, plan maintenance, drills, and similar supportive data that usually are not needed at the time of an emergency. Members of the EMO and others with assigned response duties should be able to quickly locate and refer to the appropriate action guidelines, to determine that key functions are being performed.

Using this approach, the outline of an ERP for a multi-process chemical facility would look like this:

Section I
- A. Table of Contents
- B. EMO Structure and Tasks
- C. Notification: On-Site
- D. Accounting for Personnel
- E. Headcount Center
- F. EMO Center
- G. Emergency Scene
- H. Notification: Off-site

Section II
- A. Table of Contents
- B. Work Unit 1
  1. Emergency types
  2. Action checklists
  3. Responsibilities
  4. Communications
  5. Shutdown procedures
- C. Work Unit 2 (same as above)

Section III
- A. Table of Contents
- B. Plan Philosophy and Objectives
- C. Training Programs
- D. Drills and Tests
- E. Return to Normal
- F. Appendices
  1. Special situations (e.g., bomb threats)
  2. Hazardous material data
  3. (Others as required)

Emergency Management Organization Structure

Perhaps the most essential element of an ERP is the establishment of the organization that will manage emergency response once the event has developed beyond a Level One Emergency. The EMO must be capable of rapid assembly, and must be able to carry out all of its assigned responsibilities.

The EMO is composed of the facility manager, the emergency response coordinator, the field coordinator, and various service chiefs. The number of service chiefs and their duties will vary. There need not be a separate individual in charge of each service, and small facilities may combine two or more services under one individual, as long as each function is performed.

The facility manager, in cooperation with the emergency preparedness team, assigns people, by their title or function at the facility, to the EMO, with designated alternates for each EMO function. The duties and responsibilities of each EMO function are defined and documented in the ERP. Because there must be a functional EMO in place regardless of when a facility emergency occurs, designated shift personnel take specific EMO responsibilities until the primary members can reach the site.

All of the positions established for the EMO should have written position descriptions, which may be placed in the third section of the ERP as an appendix to the plan. In the first section of the ERP, the EMO summary need only include a concise checklist for each position, with diagrams showing relationships and major communication flow paths.

Facility Manager. In the EMO, the facility manager has the overall responsibility for protecting people on-site, facility property, the environment, and the public during and after an emergency. With the assistance of the public affairs and communications chiefs, the manager usually serves as the spokesperson for the facility and the company, communicating with representatives of the media and other concerned audiences.

Emergency Response Coordinator. The designated emergency response coordinator, who may also serve as the chairperson of the emergency preparedness team, directs all emergency control activities from the MECC. All other service chiefs on the EMO report to the coordinator. Using information from the emergency scene and from the service chiefs, the coordinator makes the key decisions on what should be done, and coordinates activities on and off the site. The coordinator reports to the manager, who should be available nearby for overall guidance and counsel.

Field Coordinator. The job of the field coordinator is to correct the emergency situation as rapidly as possible with minimum risk to those in or near the affected area. He or she establishes a field command post as close to the scene of the emergency as can be safely done. The post often is in or near a radio-equipped building or vehicle, thus permitting rapid establishment of communications between the command post and the emergency response coordinator.

Ideally, the field coordinator is familiar with the operations and materials involved in the emergency; so she or he often is the highest-ranking supervisor of the affected area who is available at the time of the event. That supervisor may remain as field coordinator for the duration of the emergency, or may be relieved by another designated member of management. In the latter case, the area supervisor may become a service chief with greater hands-on involvement, with the field coordinator concentrating on marshaling required resources and maintaining effective communications. The choice depends to a great extent on the size and complexity of the facility, and the resources available for the EMO.

Public Affairs Chief. Working in close coordination with the facility manager, the public affairs chief releases appropriate information to the news media, regulatory personnel, government officials, and other public groups and individuals. No information

is to be made public by anyone other than the facility manager without specific direction from the public affairs chief. The objective is to provide full and accurate statements in a timely fashion, so that public attention is focused on facts and useful information rather than on rumors and speculation.

The public affairs chief also establishes and monitors the media headquarters, which is isolated from the EMCC to avoid interference with operations there. The public affairs chief's function includes assisting the manager in the preparation of formal statements and background information to be distributed to reporters. It also may include arranging with local radio and television stations to make periodic announcements during an emergency so that the public and the employees not on the site are properly informed. The public affairs chief also arranges to monitor and perhaps record the broadcasts of local radio and television stations in order to determine what additional statements or clarifications to the public may be required.

Communications Chief. The communications chief establishes and maintains communications capability with appropriate people on and off the site. This chief must be familiar with the various communicating systems available, including telephones, public address systems, two-way radios, and messengers. The job includes recommending and coordinating revisions or additions to the communicating systems during an emergency, and assisting the facility manager and the public affairs chief with any communications as needed.

Engineering Chief. The engineering chief's primary responsibility is to maintain electrical power for vital services. These services include on-the-scene portable lighting, continuous fire pump operation, and a steady supply of nitrogen, steam, and process cooling water. Another key duty is to assemble repair groups capable of isolating damaged sections of pipelines, electrical lines, and other necessary equipment in order to contain problems and maintain or restore operations outside the affected area. Because these repair groups must be drawn largely from site personnel such as electricians, welders, riggers, and pipefitters, the position of engineering chief should be filled by someone from the facility maintenance organization.

Rescue Chief. The rescue chief's immediate concern is to direct the rescue of trapped and/or injured people on the site. He or she identifies any outside resources needed to conduct the rescue, and coordinates their efforts when they are on the facility. The work requires good communications and coordination with the field coordinator and other chiefs, particularly the engineering chief, because rescue may involve material handling and temporary construction skills. For this reason, the rescue function is sometimes part of the engineering chief's responsibilities.

Emergency and Fire Chief. The emergency and fire chief is responsible for fire-fighting and fume control activities. The best person for this job has good knowledge of and access to the fire-fighting and fume control resources available on and off the site. On a small facility this may be the captain of the fire brigade. The job includes direction of designated facility personnel, and coordination with any outside forces brought in to bring the fire or fume condition under control. The environmental chief may assist in determining optimum courses of action based upon actual or threatened adverse effects on air and water emissions from the facility.

Medical Chief. The medical chief ensures that the proper medical care is provided to people on the facility who have been injured or exposed to toxic materials. When facilities have medical professionals on the site, one of them carries out this function. This chief's responsibilities include establishing field stations to treat affected personnel, and, in cooperation with the transportation chief, providing transportation for injured people to other medical facilities.

The medical chief also participates in discussions with community officials and

appropriate facility personnel regarding actual or potential medical problems for people outside the site boundaries. Representatives from nearby local hospitals, ambulance services, fire departments, police forces, and emergency management groups also may be involved in these discussions. The medical chief communicates any action or contingency plans developed in the discussions to the emergency response coordinator and the public affairs chief.

Environmental Chief. The environmental chief oversees all activities designed to minimize adverse effects upon the quality of air and water as a result of an emergency. The function includes coordinating air and water quality monitoring on and off the site during and after the emergency, and assisting the emergency and fire chief in selecting the optimum approach to abating a fire or fume condition. The environmental chief also provides assistance in projecting the path and concentration of a fume release, using computer modeling (if available) or maps with plastic overlays and developing an effective plan of action.

Personnel Chief. The personnel chief is responsible for accounting for all personnel on the facility at the time of the emergency. Unit emergency programs must include training on how this is accomplished. Unit supervisors initiate the process with headcounts at designated rally points. The success of rescue and medical activities depends on how quickly and thoroughly this information is obtained and reported to the personnel chief.

The personnel chief also coordinates communication with relatives of injured or deceased employees, and makes certain that this is completed before any names are released publicly. This requires close coordination with the public affairs chief and the manager.

The personnel chief need not be located at the MECC, but it is imperative that the personnel chief be in close touch with other members of the EMO. Telephones in the headcount center should have answering and message recording equipment to capture any unit reports arriving before the headquarters is staffed.

Protection Chief. The protection chief makes sure that entry to and egress from the facility are properly controlled. This involves securing gates; limiting entry to authorized personnel; registering all who pass through gates; meeting visitors, including representatives of the media, and escorting them to proper locations; and controlling all traffic on the site. The function also includes communicating with local police so that access to the facility is maintained, and crowd control procedures are enforced. Usually, the individual serving as protection chief at the time of the emergency also is responsible for initiating procedures to summon key facility personnel and urgently needed outside agencies.

Transportation Chief. The transportation chief coordinates and controls all transportation on the facility. This includes directing the assembly of available vehicles and crews, and identifying needs beyond site capabilities such as cranes, trackmobiles, and bulldozers that must be obtained from outside organizations. The function also includes providing suitable transportation for facility employees who monitor the effects of emergencies beyond site boundaries, or interact with community officials at an off-site location.

TRAINING PERSONNEL

Having a written facility ERP that is supported by established work unit emergency plans and procedures is an important part of the manager's job of ensuring that unplanned events will be promptly controlled with minimum risk to people, property, and the environment. There is much important work to be done, however, before the ERP is anything more than a valuable resource. Facility personnel must be trained to use the ERP effectively. There must be frequent drills

to test the plan and the people against the standards established by management as well as those established by law.

Within the facility, there are three groups of people who require training. First, there are members of the fire brigade, who must be trained to fight fires at the level specified by facility management. Next are the employees who have been assigned active roles in controlling emergencies of all types. This group includes members of the EMO and their alternates, plus designated support people such as headcount coordinators, guards, and emergency repair personnel. Finally, everyone else on the facility must have a basic understanding of the ERP, and must know how to respond when specific alarms sound.

The facility manager is responsible for ensuring that the appropriate training and retraining are done in a timely and effective manner. An employee should be trained when he or she is hired, at least annually thereafter, and when the employee's work area changes or the plan is revised. Most managers make this primarily a line organization function. They expect facility supervisors to use all available resources and means of education to accomplish the tasks. Some specialized training assistance, however, must be provided, particularly for the members of a fire brigade.

Fire Brigade Training

Training programs for fire brigades have two major objectives. One is to inform the brigade members of new hazards at their facility and innovations in fire-fighting techniques and equipment. The other objective is to provide hands-on training for developing skills in emergency operations and using equipment, including:

- Portable fire extinguishers
- Hoses and accessories
- Portable lighting
- Forcible entry tools
- Ladders
- Salvage equipment
- First-aid supplies
- Replacement parts
- Personal protective equipment
- Transportation equipment

A comprehensive training program for fire brigades must include classroom and hands-on training.

Outside resources can provide valuable assistance in the education and training of brigade members. Local fire departments and state fire schools are usually enthusiastic partners in such efforts. Often they are the key to securing adequate resources at a reasonable cost. In industrialized communities, mutual aid agreements may include cooperative training provisions with other chemical facilities, which provide opportunities for even more effective use of available resources. All mutual aid agreements should require cross-training in special hazards at the other facilities.

EMO Training

Training for personnel assigned to the facility EMO can be led by the emergency response coordinator, who meets periodically with each member of the EMO to review and refine position descriptions and the associated functional checklists. An example of such a checklist is shown in Fig. 6.4. The meetings may be followed by limited drills involving only the people and responsibilities included in the individual EMO function. Some facility managers assign specific emergency response training duties to each member of the EMO. For example, the engineering chief organizes, equips, and trains the repair groups who will be called on to physically stabilize a situation at the time of an emergency. The emergency and fire chief could be given the responsibility for maintaining a trained force of fire and fume fighters, which includes competent leadership on all shifts.

Employee Training

There are many ways to be sure that all employees understand the ERP. The emergency preparedness team can assist in the

____ 1. Personnel Chief and aides report to head count room at the MECC.

____ 2. Replay audio counts already received.

____ 3. Record work unit head counts as they are received.

____ 4. Obtain copy of visitors' log and employee with area head counts.

____ 5. Monitor attempts to locate missing personnel.

____ 6. Report to Emergency Response Coordinator as personnel are confirmed as missing.

____ 7. Issue final report to Emergency Response Coordinator after all work unit head counts are received.

____ 8. Determine from Emergency Response Coordinator names of any injured people.

____ 9. Coordinate communications with families of those missing or injured.

Fig. 6.4. Personnel chief functional checklist.

assessment of existing training programs, alerting the organization to training weaknesses and suggesting or providing creative ways to overcome those deficiencies. Unit supervisors must periodically review the ERP in scheduled group safety meetings or with individuals. Key plan elements can be reproduced on wallet-size cards, desktop displays, or telephone stick-ons. Individual or group discussions of how to react in given situations can detect weaknesses in procedures, training, or understanding.

Training programs for the three groups of employees discussed here should be documented, reviewed regularly, and included in a reference section of the facility ERP. But no matter how comprehensive the training programs may be, their effectiveness is unknown until a drill of the ERP is conducted and its systems and procedures are tested.

Facility Drills

To evaluate the effectiveness and completeness of an ERP, a facility must conduct periodic announced and unannounced Level Two and Level Three emergency drills. These are in addition to the more frequent Level One unit drills that are held to ensure that the more limited response procedures of a work unit are complete and well understood. Major internal emergency drills should be held at least four times each year, and scheduled to involve each working shift at least once a year.

To gain the maximum benefit from the drills, assigned observers should witness all aspects of response activity, and gather soon after to participate in a verbal evaluation of actions taken. These observers should include members of the emergency preparedness team. This process is enhanced by capturing on-scene action on videotape or in still photographs. Critiquing can be extremely valuable in identifying necessary plan changes, training needs, and resource deficiencies of all kinds. The process of critiquing must be controlled and managed. The leader must be sure to identify the strengths as well as the weaknesses that were revealed in the drill. The objective is to stimulate actions for positive change, not to assign blame.

Drills should be held on weekends and at night occasionally to test segments of the organization that work at times when all of the specialized resources of the facility are not immediately available. Initially, limited scale drills can test segments of the ERP, such as manning the MECC with the shift personnel that are available and carrying out a headcount without involving daytime employees.

As the proficiency of the organization increases, drill scenarios can become more

complex. Complicating factors approaching worst real-life conditions should be introduced periodically, including:

- Telephone switchboard overload.
- Absence of key EMO members.
- Arrival of major TV network anchorperson.
- Simulated mass casualties.
- Two-way radio failure.
- Evacuation of primary and backup MECCs.
- Major community impact.

The emergency preparedness team, with approval of the manager, designs the drill scenarios, monitors the organization's performance, provides leadership for critiquing, and recommends corrective actions. The team also develops and monitors a plan for involving the community in the important task of integrating a facility's ERP into the public emergency preparedness programs of the region.

INVOLVING THE COMMUNITY

Of major concern to the chemical industry is the public perception that facility managers have little concern for the welfare of their neighbors. Managers themselves have contributed to such false impressions by failing to interact with their communities in a consistent and meaningful way.

The public clearly wants to know more about the risks presented by a chemical facility. Increasingly, the public wants to help decide which risks are acceptable and which are not.

Recognizing that industrial facilities exist only with the consent of their host communities, most industrial organizations are assigning a high priority to building stronger bridges with their neighbors. The chemical industry in particular is finding that the involvement of the community in the process of emergency preparedness planning presents an excellent opportunity for constructive two-way communications. This has proved particularly productive if these communications are part of an aggressive and continuous risk communications process.

Communications

Each chemical facility needs to create opportunities for sharing information with its surrounding community; it should not wait for these opportunities to occur. Facility managers in particular should be active in the process, and should monitor the performance of their organizations. Some items that should be on a manager's checklist for risk communications are shown in Fig. 6.5.

The process should start with employee communications. In its employees, a facility has an important, informal communication link with its community that is often ignored. Employees deserve to know at least as much about their facility as their neighbors, and they deserve to know it first. They should be aware through communications (if not through actual involvement) of the facility's entire emergency preparedness program, from

1. Open communications up and down the line organization.
2. Scheduled two-way communications with all employees at least twice per month.
3. Newsletter for employees, pensioners, and key community leaders.
4. Regular meetings with local media reporters and editors.
5. Facility open house at least once every three years.
6. MSDS available to appropriate organizations in the community.
7. Meetings with political leaders and activists at their locations and at the facility.
8. Leadership and participation in local emergency planning committee.
9. Leadership in community projects such as wildlife protection, public land-use designation, etc.
10. Participation in school programs for children.

Fig. 6.5. Risk communication checklist.

prevention to preparation. The employees also should be familiar with the products made and their end uses. They should understand the potential hazards of the processes and materials with which they work, and how to protect themselves and the public from those hazards.

Other communication channels should be developed and regularly used. Scheduled meetings with representatives of community emergency service groups are useful for exchanging information and objectives concerning emergency preparedness, for promoting the sharing of resources, for gaining familiarity with one another's physical facilities and people, for identifying problems, and for recommending action for their solution.

In the 1980s, some chemical facilities and their communities formalized this approach and broadened participation in their meetings to include public officials and representatives of regulatory agencies and the media. Operating as Hazardous Material Advisory Councils (HMACs) and meeting regularly with established leadership and agendas, these groups quickly became key resources in organizing a community's efforts to better understand the potential for hazardous material incidents and to protect against them. Typically the responsibilities of an HMAC include coordinating a regional risk assessment, assisting the development of a community response plan specific to hazardous materials, and assisting with educational programs for various segments of the public.

The Responsible Care program of the Chemical Manufacturers Association embraces HMACs as one good way to communicate relevant and useful information that responds to public concerns for safety, health, and the environment. However, managers are finding many other ways to interact with the public to achieve a fuller measure of community awareness and involvement in affairs of mutual interest. Among the many options from which a manager may choose are the following:

- Hosting facility tours featuring emergency prevention and mitigation procedures.
- Speaking at community meetings (service clubs, schools, governing bodies, etc.).
- Sending newsletters to selected neighbors.
- Preparing informative brochures or newspaper inserts.
- Appearing on local TV or radio.

An open and sincere comprehensive risk communication process led by the facility manager creates a better-informed public that is able to understand real risks (vs. perceptions) and is likely to respond effectively in case of an actual emergency.

Integrating Plans

Most communities have long had written emergency response plans designed for natural events such as floods and windstorms; some communities have had written plans dealing with emergencies created by people, such as bomb threats and civil disturbances; but, until recently, few had specific plans for responding to emergencies involving hazardous materials. As a result, the consequences of accidental chemical releases have been in many cases tragically magnified by the undisciplined reactions of people near the release source. It has been reported, for example, that when the alarms sounded at Bhopal, residents of the nearby homes ran toward the plant rather than taking action to protect themselves from the enveloping fumes.

It is not enough to train the personnel of a chemical facility to implement an ERP effectively. Appropriate people in the community, especially near neighbors, need to understand the elements of an ERP that are designed to protect them and the role they play in making the plan work. There must be a continuous effort to integrate the facility ERP into community emergency planning at local and district levels. Drills involving external resources that test all of the plan elements against the standards mutually established with the community and those imposed by laws and regulations are necessary to ensure successful implementation of the plan.

Off-site Warning

Designing an effective off-site warning system presents some major challenges. Despite excellent ongoing communications between a chemical facility and its neighbors, there is no positive way to ensure that the general public will respond quickly and appropriately to a warning alarm of any kind. Thus, it is essential that the selected warning system be developed with the close cooperation of the community. Even then, it is difficult to predict such factors as the inclination of people to be warned and the degree of public confidence in the validity of an alarm.

No off-site warning system will assure complete coverage of the intended audience. Best results are achieved by combining two or more systems for sequential alerting—the first to trigger preplanned immediate action by the public at greatest risk, followed by other communications that provide further information and guidance to a larger audience. Some of the systems most commonly used are:

- Facility fixed-sound sources, such as sirens and whistles.
- Mobile alerting by police or fire personnel, either from vehicles with loudspeakers or door-to-door.
- Fixed public address systems in the community or in the facility.
- Automated telephone calling.
- Alert radios energized by a special signal to produce a warning tone followed by broadcast messages.
- Strobe lights in situations where the noise level is a problem.
- Local radio stations and the emergency broadcast system.
- Local TV stations.

More sophisticated and less commonly used warning systems include helicopters equipped with loudspeakers, modified cable TV installations, and computer networks between a chemical facility and community emergency response groups.

Local Emergency Plans

Existing plans for a coordinated response to emergencies in a community vary greatly in content and organization, but the plans have two common objectives. They are to:

- Define authority and responsibilities of various emergency service participants.
- Describe the interaction between those participants, government, and industry.

In creating their plans, most communities draw on the Integrated Emergency Management System (IEMS) developed by the Federal Emergency Management Agency (FEMA).

A local plan has many of the same elements as a chemical facility ERP. It includes:

- An emergency management organization, with designated functional responsibilities.
- The location of the emergency operating center and its resources.
- Guidelines for classifying emergencies.
- Activation and declaration checklists.
- Communications requirements and available systems.
- Evacuation and sheltering plans.
- Methods for securing added resources.
- Descriptions of local hazards.

Most local plans are written to be nonspecific as to the cause of the emergency, with various appendices describing the details of response to specific events. These appendices are based upon the results of risk assessments made by the community with the cooperation of industry.

One such appendix should relate to emergencies caused by fixed facility or transportation incidents in which hazardous materials are involved. Chemical facilities must provide substantial support to the community in preparing this portion of the local emergency plan, and provide resources and training leadership that are not available elsewhere in the community. Where a Hazardous Material Advisory Council exists, there is an effective

forum for doing this. In any case, a chemical facility manager should seek ways to help the community prepare for and recover from incidents of this nature.

Local Emergency Planning Committees

An important contribution to community and industry cooperation in emergency preparedness was the passage in 1986 of the Superfund Amendments and Reauthorization Act (SARA), which contained an emergency planning and community right-to-know provision. Title III, or EPCRA, as this portion of SARA is commonly called, is intended to encourage and support hazardous materials emergency planning efforts at the state and local level, and to provide citizens and local agencies with appropriate information concerning potential hazards in their communities.

The major portions of Title III require:

- A statewide organization for planning emergency action and receiving hazardous chemical information.
- Notification to the community of emergency releases of chemicals.
- Reports of hazardous chemical inventories and copies of MSDSs to be furnished to the community.
- An annual inventory of hazardous chemical releases to the environment.

Drills and Critiques

The optimum frequency of major drills involving personnel outside a chemical facility is dependent upon a number of variables:

- Location of the facility.
- Dependence upon community emergency agencies.
- Size and complexity of the facility.
- Site and off-site risk assessments.
- Population patterns.

An important element of emergency preparedness is the establishment of an appropriate major drill frequency in cooperation with off-site agencies. A reasonable goal is to hold one such drill each year. The scenario might include an on-site, internally generated hazardous material emergency one year and a transportation emergency somewhere in the adjacent community the following year.

For facilities and communities just beginning to test their plans, desktop or simulated drills are effective for identifying procedural problems that need to be corrected before they proceed to full-scale drills. In these simulations, staffing of the appropriate emergency center would occur, but the emergency service groups would not actually mobilize at the scene of the incident.

As people gain confidence in the completeness and the effectiveness of the emergency response plan, it becomes important to measure the performance of all who are involved. Monitors record and later report on all aspects of response actions, including:

- Elapsed times before critical actions occur.
- Actions and coordination of responding groups.
- Actions of uninvolved personnel.
- Alarm and communication effectiveness.
- Emergency control center management.
- Control at the emergency scene.
- Accounting for personnel.
- Medical aid for simulated casualties.
- Off-site notifications.
- Handling media representatives—real or simulated.

Following each drill there must be an organized critique that provides the information needed to strengthen the plan and/or its implementation. All of the people actively involved should be represented at the critique, and a written report of conclusions and recommendations should be widely distributed. It then is the responsibility of the facility emergency preparedness team and the local emergency planning committee to coordinate and assist in solving any problems identified—a process that begins emergency

preparedness activities again: identifying hazards; evaluating and strengthening resources; modifying the emergency plan; training people; communicating and integrating plans; and testing them once again.

LAWS, REGULATIONS, AND SUPPORT

Laws

A number of legal requirements must be incorporated in a facility's emergency response program. Emergency prevention, preparedness, and response planning are regulated at the federal, state, and, occasionally, local levels. At the federal level, these laws include

- Clean Air Act (CAA)
- Clean Water Act (CWA)
- Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA)
- Emergency Planning and Community Right-to-Know Act (EPCRA, or SARA Title III)
- Energy Reorganization Act (Nuclear Regulatory Commission)
- Hazardous Materials Transportation Act (HMTA)
- Occupational Safety and Health Act (OSHA)
- Resource Conservation and Recovery Act (RCRA)
- Toxic Substances Control Act (TSCA)
- U.S. Coast Guard requirements

A list of these laws and their *Code of Federal Regulations* (CFR) citations appears in the bibliography at the end of the chapter. States may have their own laws and regulations that also govern emergency response planning. References to these laws may be found in the Bureau of National Affairs (BNA) *Environment Reporter*.

These laws and their regulations are enforced by all levels of governmental agencies. A knowing or willful violation has serious implications for both companies and individuals, who may be held civilly or criminally liable for noncompliance. The penalties can be severe, ranging from daily-assessed fines to imprisonment. Thus, it is prudent to know the regulations that apply to the facility, and to ensure that the regulatory requirements are met.

The regulations regarding emergency planning and response are comprehensive, covering every aspect from prevention to reporting. RCRA's regulations cover the entire process, from planning to training to formal reports on the facility's response to an emergency involving hazardous waste. Other regulations focus on a specific aspect of emergency response, or part of the facility's operations. For example, TSCA requires that spills or releases that contaminate the environment be reported orally and in writing within a certain time frame. All but one of the laws put the burden of planning on the facility. EPCRA, the exception, requires state and local agencies to prepare an emergency response plan for the community. Facilities that meet criteria specified in EPCRA regulations have to assist in the development of the plan; however, they are not responsible for creating it.

Inevitably, some of the regulations and requirements overlap. Most chemical facilities are subject to more than one law, and could be expected to prepare separate plans for specific parts of their facility. RCRA's Contingency Plan, for example, must be developed and maintained apart from other emergency response plans. The key to managing all of the requirements and satisfying the regulations in an efficient, coordinated manner is first to understand the requirements and how they apply to the facility, and then to look for the common denominators among the requirements. The finished product, or master plan, will satisfy all of the common denominators that apply, and will avoid duplication of effort. It also can be used as the basis for plans that must be maintained separately or that have requirements in addition to the common denominators.

Meeting the Requirements

Regulations governing emergency response planning can be broken down into four

general categories:

- Preparation
- Plans
- Reports and other communications
- Drills and evaluations

Figure 6.6 shows a matrix-type summary of the major federal laws and their requirements for emergency planning and response. Such a matrix is very helpful in determining what the requirements are and how they apply to a facility. It could be further tailored to cover only the requirements that apply to a specific site or operation.

Prevention and Preparation

Some regulations require that a facility conduct a risk assessment and/or other preparatory activities. RCRA calls it a preparedness and prevention plan. A facility subject to RCRA must determine how structures, processes, and operations can be changed in order to minimize the possibility of an emergency involving hazardous waste. The facility also has to determine the communications and alarm systems that will be used in the event of such an emergency. CWA includes prevention in its requirements for the Spill Prevention, Containment, and Countermeasure (SPCC) Plan. The 1990 amendments to the CAA added an accident prevention plan for extremely hazardous substances.

Plans

At the very least, a facility is required to develop a plan describing how it will respond to an incident that threatens human health and/or the environment. Generally, the plan includes notification, evacuation, protection of employees, and control of the incident. This emergency response plan usually must be in writing. For example, OSHA requires a minimum of three plans: emergency response, emergency action, and fire prevention. The CAA requires that the state implementation plan have an emergency air pollution episode plan.

Communications

There are two aspects to emergency communications: the actual equipment used to communicate information about the incident and the types of communications or information sharing required. RCRA has specific requirements for the types of emergency communication equipment (alarm systems, phone or radio communications) that must be present. Under EPCRA, facilities must provide information about their operations and substances used or stored on site when the Local Emergency Planning Committee (LEPC) or State Emergency Response Commission (SERC) requests it. If the facility uses or stores extremely hazardous substances (EHS) in reportable quantities, it must appoint a representative to the LEPC. Several laws require that a copy of the emergency response plan be made available to employees and representatives of government agencies during working hours. OSHA requires that facilities provide material safety data sheets (MSDS) for all hazardous substances present on the site.

Reports

Reports are another important communications aspect of the emergency response program. Most laws insist on prompt notification of the proper agencies immediately after an incident occurs. These reporting requirements can be complex, particularly in view of the fact that many laws have their own lists of hazardous substances and reportable quantities. CERCLA requires that releases of a reportable quantity of what it defines as a hazardous substance must be reported immediately to the National Response Center (NRC). EPCRA requires facilities that store and/or release reportable quantities of substances on its EHS list to report that information to the LEPC and the SERC.

Written followup reports are often a requirement. HMTA has identified six specific

	RCRA	OSHA	DOT/HMTA	EPCRA	CERCLA
PREVENTION AND PLANNING	Preparedness & prevention plan Contingency Plan (CP), in writing ❑ Emergency Coordinator ❑ Evacuation ❑ Access ❑ Equipment ❑ Communications	Emergency Response Plan ❑ For entire site ❑ For employees who respond to uncontrolled releases of hazardous substances, including hazardous waste Emergency Action Plan ❑ Evacuation for employees in case of incidental chemical releases ❑ How to report an incident Fire Prevention Plan	Incident prevention and response	Local emergency planning committee (LEPC) prepares emergency response plan for community. State emergency planning commission (SERC) oversees.	National Contingency Plan (NCP) Facility plan for response and cleanup of oil or hazardous substance must meet NCP standards.
REPORTS	Incidents in transit Transporter must ❑ Notify National Response Center (NRC) at once. ❑ Submit written report within 15 days to DOT. ❑ Coordinate with DOT. Hazardous waste emergency on site ❑ Immediately-NRC ❑ Follow-up, in writing to EPA RA within 15 days ❑ UST Releases		Report six specific hazardous material incidents, at once to DOT/NRC; written follow-up at once, to NRC. Spill of RQ into navigable waters	Releases of extremely hazardous substance (EHS) over reportable quantity (RQ) EHS stored on site in quantities ≥ RQ	Releases of RQ or 1 pound of hazardous substance, to NRC immediately
INFORMATION	Maintain copy of CP at site. Provide copy to local emergency response organization. Establish alarm and communication systems for emergency notification.	MSDSs to employees and emergency response organizations	Emergency response information available during transportation and at facilities where hazardous materials are loaded or stored.	As requested by LEPC and SERC. Designate representative to LEPC. Coordinate internal plans or make them available to LEPC.	
TRAINING, DRILLS, AND EVALUATION	Emergency response must be documented, records retained. Initial and annual review Evacuation drill	Initial and annual refresher training for employees involved in emergency response. Requirements vary with roles. All employees trained in Emergency Action Plan: initially and with every change to plan. Training in MSDS information			

Fig. 6.6. Emergency prevention, planning, and response matrix.

hazardous material incidents that must be reported immediately and again in writing. RCRA gives managers 15 days after a hazardous waste emergency to submit a written report to EPA.

Training, Drills, Audits, and Evaluations

An emergency plan is relatively useless unless the employees affected by it are trained in its use. RCRA, OSHA, HMTA, CWA, and the Energy Reorganization Act require annual and refresher training. In addition, the facilities must keep records of the training, and must make them available to the appropriate agency when they are requested. Some laws go so far as to require practice drills. Nuclear power plants must conduct on- and off-site training, and go through annual graded emergency response exercises. The plan and the response executed according to the plan

CWA	CAA	TSCA	NUREG 054	USCG
Spill Prevention Control and Countermeasure (SPCC) plan. Shows how facility will ❑ prevent ❑ respond ❑ follow-up to oil spills in harmful quantities. Must be in writing.	State Implementation Plan (SIP) must include an Emergency Air Pollution Episode plan. Accident Prevention Program for EHSs States may require prevention and emergency response plans.		Emergency response plan, including ❑ Emergency Planning Zones ❑ Prevention, mitigation, and limitation of core damage and consequences of release Subject to annual review	
Release of RQ of oil or hazardous substance to NRC, immediately	Announce uncontrolled releases of pollutants over certain set levels. Releases ≥ RQ to NRC	Emergency incidents of environmental contamination: ❑ At once to EPA ❑ Written follow-up within 10 days	Notify state and local officials ❑ Change in conditions ❑ Protective action recommendations Notify community within 15 minutes, 100% notification within 15 miles	All spills into navigable waters of oil, petroleum products, or hazardous substances, ≥ RQ, at once, to U.S. Coast Guard or to NRC.
Keep copy of SPCC at site available to EPA RA during normal working hours. Submit SPCC amendments to EPA RA and state water pollution control agency.				
Train ❑ SPCC responders. ❑ Employees who operate maintain equipment.			Annual training ❑ On-site ❑ Off-site Annual graded emergency response exercises	

then are evaluated so that the plan can be improved.

Sources of Assistance

Seeing all of the requirements together can be overwhelming. Fortunately, there are agencies, associations, and programs that can assist in the preparation of a comprehensive emergency prevention and response plan.

The federal government and the agencies responsible for the laws that govern emergency response planning provide 800-number hotlines and manuals that describe various aspects of emergency prevention, planning, and response.

The volunteer or professional emergency responders in the community have valuable practical experience that can be put to work in developing the facility emergency response plan. Working with them also establishes a

forum for communications and understanding with the community.

The Chemical Manufacturers Association (CMA) Community Awareness and Emergency Response (CAER) program provides comprehensive guidelines for the development and implementation of an emergency response plan. The CAER program has been expanded to include all aspects of the chemical industry in an initiative called Responsible Care.

Other services of the CMA include CHEMTREC, a 24-hour emergency response service for people who respond to emergencies involving chemicals; CHEMNET, a mutual aid agreement between chemical producers and emergency response contractors; and workshops and videotape training programs for first responders and other emergency response personnel.

Additional sources of assistance and information include other professional associations, such as the American Institute of Chemical Engineers, and publications, seminars, workshops, and videotapes offered by educational organizations.

A bibliography; a list of laws, regulations, and standards; and a compilation of suggested reading material follow.

SELECT BIBLIOGRAPHY

Chemical Manufacturers Association. *CAER: The Next Phase Program Handbook.* Washington, DC: Chemical Manufacturers Association with assistance from HMM Associates, Concord, MA, 1989.

———. *Community Emergency Response Exercise Handbook.* Washington, DC: Chemical Manufacturers Association, 1986.

———. *Emergency Warning Systems Guidebook.* Washington, DC: Chemical Manufacturers Association, 1987.

———. *Site Emergency Response Planning Handbook.* Washington, DC: Chemical Manufacturers Association, 1986.

The Conservation Foundation. *Risk Communication: Proceedings of the National Conference on Risk Communication.* Washington, DC: The Conservation Foundation, 1986.

Covello, V. T., Sandman, P. M., and Slovic, P., *Risk Communication, Risk Statistics, and Risk Comparisons.* Washington, DC: Chemical Manufacturers Association, 1988.

Du Pont Company. *Du Pont Safety and Occupational Health Guidelines.* Wilmington, DE: E. I. du Pont de Nemours and Company, 1990.

———. *Major Incident Preparedness.* Wilmington, DE: E. I. du Pont de Nemours and Company, 1987.

———. *Title III, Volume I, Guidelines: Emergency Planning and Right-to-Know.* Wilmington, DE: E. I. du Pont de Nemours and Company, 1989.

Federal Emergency Management Agency. *FEMA Handbook.* Washington, DC: Federal Emergency Management Agency, 1985.

National Response Team. *Hazardous Materials Emergency Planning Guide (NRT 1).* Washington, DC: National Response Team of the National Oil and Hazardous Substances Contingency Plan, 1987.

———. *Criteria for Review of Hazardous Materials Emergency Plans.* Washington, DC: National Response Team, 1988.

Occupational Safety and Health Administration. *How to Prepare for Workplace Emergencies, OSHA 3088 (revised).* Washington, DC: U.S. Department of Labor, 1988.

Prugh, R. W., and Johnson, R. W., *Guidelines for Vapor Release Mitigation.* New York: Center for Chemical Process Safety, The American Institute of Chemical Engineers, 1988.

U.S. Environmental Protection Agency. *Guide to Exercises in Chemical Emergency Preparedness Programs.* Washington, DC: U.S. Environmental Protection Agency, 1988.

LAWS, REGULATIONS, AND STANDARDS

Laws and Regulations

CAA (Clean Air Act)
- 40CFR 51 SIP Emergency air pollution episode plan
- CAA Reauthorization Section 304, Prevention of sudden, accidental releases

CERCLA (Comprehensive Environmental Response, Compensation, and Liability Act)
- 40CFR 300-306 Section 103(a) Spill reporting requirements
- 40CFR 355 Emergency planning

CWA (Clean Water Act)
- 40CFR 112.3-7 Spill prevention, control, and countermeasure plan (SPCC) requirements, modifications, preparation, and implementation

Energy Reorganization Act (was Atomic Energy Act)
- 10CFR 50.47, also Appendix E NRC (Nuclear Regulatory Commission) Standards for on-site and off-site emergency plans

HMTA (Hazardous Materials Transportation Act)
- 49CFR Parts 171-177 DOT requirements for providing information and advice on meeting emergencies; FEMA requirements for evaluation of training programs for incident prevention and response

OSHA (Occupational Safety and Health Act)
- 29CFR 1910.1200 Hazard communication standard
- 29CFR 1910-210 and Appendices, Hazardous waste operations and emergency response
- 29CFR 1910.35-38 Requirements for evacuation route and exit posting, emergency lighting, accessibility of exits, and emergency action plan
- 29CFR 1910.156 Fire brigades
- 29CFR 1910.157-164 Fire extinguishing and detection systems
- 29CFR 1910.165 Employee alarm systems

RCRA (Resource Conservation and Recovery Act)
- 40CFR 263.30,31 Transporter responsibilities in hazardous waste transportation incidents
- 40CFR 264.30-37 and 40CFR 265.30-37 Preparedness and prevention
- 40CFR 264.50-56 and 40CFR 265.50-56 Contingency Plan

SARA Title III (Superfund Amendment and Reauthorization Act, Emergency Planning and Community Right-to-Know)
- 40CFR 355.3, Section 302(c), Emergency planning notification
  Section 303(d), Appointment of emergency coordinator, provision of information; Section 304, Emergency release notification requirements

U.S. Coast Guard
- 33CFR 126.9 Reporting requirements for discharge of petroleum products or dangerous liquid commodities into navigable waters of USA
- 33CFR 153.023 Reporting requirements for discharge of reportable quantity (RQ) of oil or hazardous substance into navigable waters

Standards

ANSI/National Fire Protection Association Standards
- #110 Emergency Power and Standby Systems
- #600 Private Fire Brigades

SUGGESTED READING

Regulations

Bureau of National Affairs. *Environment Reporter.* Washington, DC: Bureau of National Affairs, Inc.

ENSR Consulting and Engineering. *Air Quality Handbook: A Guide to Permitting and Compliance under the Clean Air Act and Air Toxics Programs, 10th edition.* Acton, MA: ENSR Consulting and Engineering, 1988.

ERT, Inc., and Sidley & Austin. *Superfund Handbook, 2nd edition.* Concord, MA: ERT, Inc. and Chicago, IL: Sidley & Austin, 1987.

Lowry, G. G., and Robert, C., *Lowry's Handbook of Right-to-Know and Emergency Planning.* Chelsea, MI: Lewis Publishers, Inc., 1988.

Office of Solid Waste, U.S. Environmental Protection Agency. *Solving the Hazardous Waste Problem: EPA's RCRA Program (EPA/530-SW-86-037).* Washington, DC: U.S. Environmental Protection Agency, 1986.

Prevention and Planning

Abrams, M. J., and Lewis, J., "Preplanning, the Key Emergency Response," *Papers Presented at the Spring 1987 National Meeting of the American Institute of Chemical Engineers.* New York: American Institute of Chemical Engineers, 1987.

Adler, V., Sorenson, J. H., and Rogers, G. O., "Chemical and Nuclear Emergencies: Interchanging Lessons Learned from Planning and Accident Experience." *Proceedings of a Joint NEA/CEC Workshop on Technical Aspects of Emergency Planning, Brussels, Belgium (CONF-89906137-1).* Washington, DC: U.S. Department of Energy, 1989.

American Nuclear Society. *Emergency Response—Planning, Technologies, and Implementation; Proceedings of the ANS Topical Meeting (CONF-880913, UC-610).* Aiken, SC: Savannah River Laboratory, E. I. du Pont de Nemours and Company, 1988.

Bell, D. W., and Burns, C. C., "Offsite Emergency Plan Development and Maintenance Considerations," *Papers Presented at the Spring 1987 National Meeting of the American Institute of Chemical Engineers.* New York: American Institute of Chemical Engineers, 1987.

Davis, D. S., et al. *Prevention Reference Manual: Overviews on Preventing and Controlling Accidental Releases of Selected Toxic Chemicals (EPA/600/8-88-074).* Research Triangle Park, NC: U.S. Environmental Protection Agency, Air and Energy Engineering Research Lab., 1988.

Dickerson, M. H., "Emergency Planning, response, and Assessment: A Concept for a Center of Excellence," *Proceedings from an International Seminar on Nuclear War International Cooperation: The Alternatives (CONF-8608149-1).* Washington, DC: U.S. Department of Energy, 1986.

Federal Emergency Management Agency. *Hazardous Materials Exercise Evaluation Methodology (HM-EEM) and Manual.* Washington, DC: Federal Emergency Management Agency, 1989.

Fingleton, D. J., Tanzman, E. Z., and Bertram, K. M. "Development of a Model Emergency Response Plan for Catastrophic Releases of Toxic Gases," *Proceedings of the Air Pollution Control Association Annual Meeting (CONF-860606-19).* Washington, DC: U.S. Department of Energy, 1986.

Gudiksen, P., et al. *Emergency Response Planning for Potential Accidental Liquid Chlorine Releases.* Washington, DC: U.S. Department of Energy, 1986.

International Association of Fire Chiefs. *Fire Service Emergency Management Handbook.* Washington, DC: Federal Emergency Management Agency, 1985.

Jones, E. "Contingency Planning and Emergency Response in Construction Activities: Training the Construction Worker," *Proceedings of an Oak Ridge Model Conference on Waste Problems (CONF-871075-5).* Washington, DC: U.S. Department of Energy, 1987.

Kalnins, R. V., "Emergency Preparedness and Response," *Symposium on the Characterization of Thermodynamic and Transport Properties of Polymer Systems.* New York: American Institute of Chemical Engineers, 1986.

Krimsky, S., and Plough, A., *Environmental Hazards—Communicating Risks as a Social Process.* Medford, MA: Auburn House Publishing Company, 1988.

Michael, E. J., et al. "Emergency Planning Considerations for Specialty Chemical Plants," *American Institute of Chemical Engineers 1986 Summer National Meeting.* New York: American Institute of Chemical Engineers, 1986.

Morentz, J. W., and Griffith, D., "Using Computers for Chemical Emergency Planning and Response," *Papers Presented at the Spring 1987 National Meeting of the American Institute of Chemical Engineers.* New York: American Institute of Chemical Engineers, 1987.

National Response Team. *Criteria for Review of Hazardous Materials Emergency Plans (NRT-1A).* Washington, DC: National Response Team of the National Oil and Hazardous Substances Contingency Plan, 1988.

———. *HAZMAT Planning Guide (WH-562A).* Washington, DC: National Response Team, 1987.

———. *Technical Guidance for Hazards Analysis: Emergency Planning for Extremely Hazardous Substances (Supplement to Hazardous Materials Emergency Planning Guide NRT 1).* Washington, DC: National Response Team, 1987.

Philley, J. O., "Emergency Preparedness Training Tips," *American Institute of Chemical Engineers National Meeting—Summer '87.* New York: American Institute of Chemical Engineers, 1987.

U.S. Environmental Protection Agency. *Bibliography on Chemical Emergency Preparedness and Prevention.* Washington, DC: U.S. Environmental Protection Agency, 1986.

———. *It's Not Over in October: A Guide for Local Emergency Planning Committees.* Washington, DC: U.S. Environmental Protection Agency, 1988.

———. *Seven Cardinal Rules of Risk Communication (OPA 87-020).* Washington, DC: U.S. Environmental Protection Agency, 1988.

———. "Why Accidents Occur: Insights from the Accidental Release Information Program," *Chemical Accident Prevention Bulletin (OSWER-89-008.1)*, Series 8, No. 1, (July 1989). Washington, DC: U.S. Environmental Protection Agency, 1989.

Waldo, A. B. (Ed.). *The Community Right-to-Know Handbook.* Washington, DC: Thompson Publishing Group, 1986.

Transportation

Office of Hazardous Materials Transportation, U.S. Department of Transportation. *Emergency Response Guidebook: Guidebook for Initial Response to Hazardous Materials Incidents (DOT P5800.4).* Washington, DC: U.S. Department of Transportation, Office of hazardous Materials Transportation, Research and Special Programs Administration, 1990.

7

Wood and Wood Products

Raymond A. Young*

Wood has been utilized by humans since antiquity. Trees provided a source of many products required by early humans such as food, medicine, fuel, and tools. For example, the bark of the willow tree, when chewed, was used as a painkiller in early Greece and was the precursor of the present-day aspirin. Wood served as the primary fuel in the United States until about the turn of the century, and even today over one-half of the wood now harvested in the world is used for heating fuel.

Throughout history different types of wood have served many purposes. The tough, strong, and durable white oak, for example, was a well-proven raw material for ships, bridges, cooperage, barn timbers, farm implements, railroad ties, fence posts, flooring, paneling, and other products. In contrast, woods such as black walnut and cherry became primarily cabinet woods. Hickory was manufactured into tough, hard, resilient tool handles. Black locust was used for barn timbers and treenails. What the early artisan learned by trial and error became the basis for intelligent decisions concerning which species was best suited to a given purpose, as well as what characteristics should be looked for in selecting trees for different applications. It was known that wood from trees grown in certain locations was stronger, more durable, and more easily worked with tools than wood from the same species grown in other locations. Modern wood quality research has substantiated that location and growth conditions significantly affect wood properties.[1]

In the United States more than 100 kinds of wood are available to the prospective user, but it is very unlikely that all are available in any one locality. About 60 native woods are of major commercial importance. Another 30 wood types are commonly imported in the form of logs, cants, lumber, and veneer for industrial uses, the building trades, and crafts.[2,3]

The importance of wood as a raw material supplying fiber, energy, and chemicals is similar in magnitude to its use as a solid material. Lumber, plywood, and reconstituted boards consume about one-half of the timber

*Department of Forestry, University of Wisconsin, Madison, Wisconsin. The author wishes to acknowledge the use of material from an earlier edition prepared by Dr. E. C. Jahn and Dr. R. W. Strauss.

Fig. 7.1. Summary of uses for wood. (*From* Introduction to Forest Science, *2nd ed., R. A. Young and R. L. Giese, Eds., copyright © 1990 John Wiley & Sons, New York; with permission of John Wiley & Sons Publisher.)*

harvest; usage for fiber, chemicals, and fuels accounts for the remaining half. Fiber production consumes about 25 percent of the total timber use. In 1990 over 55 million metric tons of pulp fiber for papermaking were derived from the forest. Although the relative value of wood as a source of energy and chemicals has varied considerably through the decades, wood continues to be an important source of specialty chemicals and renewable energy, and may be even more important in the future. The many uses for wood are summarized in Fig. 7.1. In this chapter we discuss the chemical nature of wood and provide a description of the technology for conversion of wood to pulp fibers for papermaking, to fuels for energy, and to chemicals for industry and consumers.

WOOD STRUCTURE

The wood from the many different species of trees is divided into two major categories according to the botanical classification of the trees as seed plants. The hardwoods are derived from the angiosperms (encased seeds), which are generally the deciduous broad-leaved trees; the softwoods are from the gymnosperms (naked seeds), generally characterized as cone-bearing trees with persistent needlelike foliage. The majority of the trees in the latter classification fall into the division Coniferophyta and thus are often referred to as conifers. It is important to note that the terms "hardwood" and "softwood" have no direct relation to the hardness or softness of the wood. In fact, hardwood trees such as cottonwood, aspen, and balsa have softer wood than the western white pines and true firs; and certain softwoods, such as longleaf pine and Douglas fir, produce wood that is much harder than that of basswood or yellow poplar.

Wood can be considered as a biological composite of hollow tubes of cellulose fibers held together by a lignin matrix gluing material. Liquids are transported up and down the trees through the tubular plumbing system. About 90 percent of the wood tissue of softwood trees consists of fiber tracheid cells for liquid conduction and support. The hardwood trees evolved after the softwoods and have specialized water conduction cells called vessels.

Each vessel consists of numerous drum-shaped vessel members with open ends, stacked one above the other to form a continuous vessel pipeline from root tips to the leaves. The vessels in oaks and some other broad-leaved trees are large enough to be seen with the unaided eye, that is, as round pores at the end surface and as needlelike grooves at the lateral surface (Fig. 7.2).[4]

In some woods such as oak, the vessels in the earlywood (produced early in the growing season) are much wider than those in latewood and form the distinct growth rings of these ring-porous woods (Fig. 7.3a). In the diffuse-porous maple, birch, and many other genera, the vessels are of the same size throughout the growth ring, the boundary of which is marked only by extremely thin cells (Fig. 7.3b). Because coniferous woods lack vessels, they sometimes are designated non-porous, in contrast to the porous broad-leaved trees, particularly those with relatively wide pores. Coniferous wood has growth rings because the earlywood tracheids have wider cavities and thinner cells than do the latewood tracheids; the earlywood–latewood contrast is especially sharp in southern pines.

On tree cross sections, rays of parenchyma cells resemble spokes of a wheel as they extend from the bark into the wood, in the radial direction of the circular growth rings and perpendicular to the tangent of the rings or to the tangential direction (Fig. 7.2). Radial and tangential are both transverse directions, that is, horizontal to the trunk of the standing tree.

Fibers and vessels die at an age of a few weeks, after attaining their final shape and chemical composition. Parenchyma cells live much longer, some remaining alive until sapwood becomes heartwood. One of the tasks of parenchyma cells is to convert sugars into storage starch to serve as food reserves for the tree, and to reconvert starch to sugars

Fig. 7.2. Diagrammatic wedge section of a five-year-old oak trunk. The term "transverse section" is synonymous with end surface and cross section. Both radial section and tangential section are lateral surfaces. (*Courtesy of U.S.D.A. Forest Products Laboratory.*)

Fig. 7.3. Photomicrograph of hardwood cross sections depicting (a) ring-porous and (b) diffuse porous wood. (*Courtesy of U.S.D.A. Forest Products Laboratory.*)

when they are needed for growth. The starch and sugar content of sapwood is one reason why most wood-damaging organisms prefer dead sapwood over heartwood, and why sapwood is never quite as durable as heartwood, except in living trees whose sapwood cells fight attacking organisms.

Practically all cavities of wood cells are interconnected, as mentioned above for the vessel members of hardwoods. Tiny openings in the cell walls of overlapping ends of tracheids provide vertical passageways for water rising from cell to cell in conifers. Thin walls of ray cells are perforated for the radial transport of sugars, and tangential connections provide for growing cells. Hence, wood is permeable in all three of its cardinal directions. The longitudinal orientation of most cells, however, causes liquids and gases to flow and diffuse many times faster in the longitudinal or fiber direction than in the two transverse directions. This explains why stacked firewood dries predominantly at the exposed ends. The radial orientation of rays allow liquids and gases to move radially up to twice as rapidly as tangentially.[4]

As a porous material, wood contains air-filled or water-filled cell cavities surrounded by cell walls (Fig. 7.3). The cell wall substance is remarkably uniform; it differs very little even between conifers and broad-leaves trees and has a constant density of around 1.5 g/cm^3. The density of wood thus depends on the proportion of the cell wall volume to the volume of the bulk wood.

Wood density is *the* indicator of nearly all wood properties and of wood quality. Relatively dense species are strong, as one might expect from their large cell wall mass. Wood species also vary in features not related to density, such as color, percentage and arrangement of cell types, and minor chemical constituents. Therefore, even woods of the same density are likely to differ, with the properties varying more than the densities.

The widespread use of wood is due to its many advantages, including its very high strength-to-mass ratio, good toughness, low cost and availability, workability, attractive appearance, and good thermal and electrical insulating properties. However, there also are a number of disadvantages of wood, which have limited its usage for many applications; these include its lack of plasticity and fusibility, shrinkage and swelling properties, deterioration due to fungi and insects, and variability. However, most of its negative properties can be corrected by proper treatment and modification, as described later in this chapter.

CHEMICAL COMPOSITION AND FIBER STRUCTURE

Wood is composed of three major polymeric materials: cellulose, hemicelluloses, and lignin. Cellulose is the common name used for the glucan present in wood, which constitutes about 42 percent of wood's dry weight. Cellulose is the primary component of the walls of cells making up wood fibers and is the main structural material of wood and other plants. Paper, paperboard, and other wood fiber products thus also are composed mostly of cellulose. The chemical structure of the cellulose macromolecule is shown in Fig. 7.4. In the plant the degree of polymerization (DP) of cellulose is approximately 14,000.[5]

Closely associated with cellulose in the wood structure and paper products are other polysaccharides called hemicelluloses, which often have been labeled as the matrix material

Fig. 7.4. The chemical structure of cellulose with the cellulose repeat unit in brackets.

of wood. In hardwoods the primary hemicellulose is a xylan (polymer of xylose), whereas in softwoods the primary hemicellulose is a glucomannan, although both of these polysaccharides occur to some extent in both types of wood. The DP of the hemicelluloses is much less than that of cellulose, in the range of 100 to 200.

Table 7.1 compares the chemical composition of extractive-free hardwoods and softwoods. As cellulose and the hemicelluloses are both polysaccharides, it is obvious that the polysaccharide of wood is by far the dominant component, making up approximately 70 percent of both hardwoods and softwoods. Additional polysaccharides may occur as extraneous components of wood, which are not part of the cell wall; for example, the heartwood of species of larch can contain up to 25 percent (dry weight) of arabinogalactan, a water-soluble polysaccharide that occurs only in trace quantities in other wood species.[5]

The third major component of wood shown in Table 7.1 is lignin. Although lignin also is a polymer, it has a different chemical structure from that of the polysaccharides. The monomeric units in lignin are phenolic-type compounds, but the exact chemical structure of lignin still is not known after 100 years of intensive research. The spaces between fibers in wood are almost pure lignin and are termed the middle lamella. Lignin is considered the gluing or encrusting substance of wood and adds mechanical strength or stiffness to the tree and to wood. Higher plants commonly are referred to as lignocellulosic because of the typical joint occurrence in them of lignin and cellulose.

TABLE 7.1 Chemical Composition and Fiber Length of Extractive-Free Wood

Component	*Hardwood (Red Maple) (percent)*	*Softwood (Balsam Fir) (percent)*
Cellulose	44	42
Hemicelluloses		
Xylan	25	9
Glucomannan	4	18
Lignin	25	29
Pectin, starch	2	2
Average fiber length (mm)	0.8–1.5	2.5–6.0

In the tree the cellulose polymers are laid down uniformly, the chains paralleling one another, and the long-chain molecules associate strongly through hydrogen bonds that develop between hydroxyl groups. These bonds create very strong associations between the cellulose macromolecules, producing very uniform crystalline structures known as micelles or microcrystallites (see Fig. 7.5).

The micelles also are associated in the tree to give long threadlike structures termed microfibrils (Fig. 7.5). However, the structure of the microfibrils is not completely uniform in terms of the alignment of the cellulose macromolecules, and the regions of nonuniformity between the micelles in the microfibrils are called amorphous regions. Thus the cellulose microfibril has a crystalline–amorphous character. Water molecules enter the amorphous regions and swell the microfibrils; ultimately this is the mechanism by which fibers and wood swell in moist or wet environments.

The final fiber cell wall structure is essentially layers of the microfibrils or macrofibrils aligned in several different directions, as shown in Fig. 7.6. The microfibrils that make up the wood fiber are visible under the scanning electron microscope, which has a greater magnification than the light microscope.

The entity holding the fibers together, the middle lamella, is almost pure lignin (90%), as mentioned earlier. For the cellulose fibers to be separated, the middle lamella lignin must be chemically removed, a process that also removes most of the hemicelluloses, or must be mechanically degraded to free the fibers for papermaking. A paper sheet then can be formed from the separated cellulose fibers by depositing them from a water slurry onto a wire screen. The water drains away and the fibers collapse, leaving a fiber mat that derives its main strength from reassociation of the fibers through many hydrogen bonds—the same type of bond that gives mechanical integrity to the fibers.[5]

Fig. 7.5. Detailed structure of cell walls. (A) Strand of fiber cells. (B) Cross section of fiber cells showing gross layering: a layer of primary wall and three layers of secondary wall. (C) Fragment from the middle layer of a secondary wall showing macrofibrils (white) of cellulose and interfibrillar spaces (black), which are filled with noncellulosic materials. (D) Fragment of a macrofibril showing microfibrils (white), which may be seen in electron micrographs (Fig. 7.6). The spaces among microfibrils (black) are filled with noncellulosic materials. (E) Structure of microfibrils: chainlike molecules of cellulose, which in some part parts of the microfibrils are orderly in arrangement. These parts are the micelles. (F) Fragment of a micelle showing parts of chainlike cellulose molecules arranged in a space lattice. (G) Two glucose residues connected by an oxygen atom—a fragment of a cellulose molecule. (*From K. Esau,* Anatomy of Seed Plants, *Second Edition, 1977, copyright © by John Wiley & Sons, 1977,* reprinted by permission of the copyright owner.)

Fig. 7.6. A section of wood made up of fibers and middle lamella. The structure of a fiber is given, showing the micro-fibrillar orientations in the different layers of the fiber cell wall. (*From* Introduction to Forest Science, *2nd ed., R. A. Young and R. L. Giese, Eds. Copyright © 1990 by John Wiley & Sons, Inc., New York, and reproduced by permission of the copyright owner.*)

The long fibers from softwoods (Table 7.1) usually are preferred in papermaking for products that must resist tearing, such as grocery bags, whereas the shorter hardwood fibers give improved opacity, or covering power, and printability to the final paper sheet. The type of pulping process also affects the pulp properties, as described in a later section.

PULP AND PAPER

The concept of making paper from the fibers from lignocellulosic materials—an integrated system of fiber separation (pulping) and re-forming of the fibers into a mat (papermaking)—is attributed to T'sai Lun, a court official in southeast China in A.D. 105. The first fibers were obtained from old hemp rags and ramie fishnets, but shortly thereafter, the inner bark fibers from paper mulbery trees also were utilized for papermaking. Bamboo was used as a source of fiber several centuries later. The rags were macerated into a pulp in water with a mortar and pestle; then, after dilution in a vat, the pulp was formed into a

wet mat on a bamboo frame equipped with a cloth screen to drain the free water. The mat was dried in the sun. The invention was based on the need for a writing material to replace the expensive silk and inconvenient bamboo strips. The invention was a closely guarded secret for many centuries but eventually filtered west, reaching Europe in 1151 and the United States (Pennsylvania) in 1690.[5]

The importance of the pulp and paper industry to the American economy is exemplified by the growth rate in the use of paper and paper products. New uses continually are being found for paper, and these developments together with a rising standard of living have resulted in a constant increase in the per capita consumption of paper. The industry has shown constant growth since about the middle of the nineteenth century, when processes were invented for production of fiber from wood, replacing reliance on cotton rags as a source of fiber (Fig. 7.7). The relative paper consumption by grade is shown in Fig. 7.8, and differences in the economies of developed and less developed countries are demonstrated in Fig. 7.9 by the wide disparity in per capita consumption of paper and paperboard.

Fig. 7.7. Consumption of paper in the United States since the year 1860.

Fig. 7.8. Relative consumption of paper by grade.

Fig. 7.9. Comparison of paper consumption in developed and developing countries.

Fig. 7.10. Paper production in the ten leading states.

In 1988, the United States produced a total of 69,477,000 tons of paper and paperboard in hundreds of separate mill locations. This paper and board was made from 55,530,000 tons of pulp, produced primarily from wood in over 200 pulp mills, and 15,000,000 tons of recycled waste paper.[6]

Figure 7.10 shows the leading states for paper production in 1988. Only a small amount of paper is made from rags or other fiber sources (agricultural residues); so most wood pulp is produced from those areas of the country that are heavily forested. With the exception of Maine and Wisconsin, pulp production is concentrated in the southern and northwestern sections of the United States. Georgia and Washington are the top pulp-producing states. Although a high percentage of pulp is converted into paper or board at the same plant site, a significant portion (classified as "market pulp") is sold in bales to mills in other sections of the country for subsequent manufacture into paper. Thus, New York and Michigan are relatively important in papermaking, but they rank quite low as pulp producers. However, both New York and Michigan are large users of waste paper that is reprocessed into usable fiber. With the current emphasis upon recycling, it is probable that the use of waste paper will be greatly accelerated in the near future.

Raw materials for the pulp and paper industry can be classified as fibrous and nonfibrous. Wood accounts for over 95 percent of the fibrous raw material (other than waste paper) in the United States. Cotton and linen rags, cotton linters, cereal straws, esparto, hemp, jute, flax, bagasse, and bamboo also are used and in some countries are the major source of papermaking fiber.

Wood is converted into pulp by mechanical, chemical, or semichemical processes. Sulfite and kraft (sulfate) are the common chemical processes, and neutral sulfite is the principal semichemical process (NSSC). Coniferous wood species (softwoods) are the most desirable, but the deciduous, broad-leaved species (hardwoods) have gained rapidly in their usage and constitute about 25 percent of pulpwood. Table 7.2 summarizes the conditions utilized with the various pulping processes, and Fig. 7.11 shows the relative production of pulp by the major processes.[5]

Nonfibrous raw materials include the chemicals used for the preparation of pulping liquors and bleaching solutions and the various additions to the fiber during the papermaking process. For pulping and bleaching, these raw materials include sulfur, lime, limestone, caustic soda, salt cake, soda ash, hydrogen peroxide, chlorine, sodium chlorate, and magnesium hydroxide. For papermaking, they include rosin, starch, alum, kaolin clay, titanium dioxide, dyestuffs, and numerous other specialty chemicals.

Wood Preparation

The bark of trees contains relatively little fiber and much strongly colored nonfibrous material; what fiber there is, is of poor quality. The nonfibrous material usually will appear as dark-colored dirt specks in the finished

TABLE 7.2 Wood Pulping by Process and Yield

Process	Acronym	Treatment: Chemical	Treatment: Mechanical	Pulp Yield (percent)
Mechanical Processes				
Stone groundwood	SGW	None	Grinder	93–95
Pressure groundwood	PGW	None	Grinder	93–95
Refiner mechanical	RMP	None	Disk refiner (pressure)	93–95
Thermomechanical	TMP	Steam	Disk refiner (pressure)	80–90
Chemithermomechanical	CTMP	Sodium sulfite or sodium hydroxide[a]	Disk refiner (pressure)	80–90 80–90
Chemimechanical[b]	CMP	Sodium sulfite or sodium hydroxide	Disk refiner	80–90
Semichemical Processes				
Neutral sulfite	NSSC	Sodium sulfite + sodium carbonate	Disk refiner	70–85
Green liquor	GLSC	Sodium hydroxide + sodium carbonate	Disk refiner	70–85
Nonsulfur	—	Sodium carbonate + sodium hydroxide	Disk refiner	70–85
Chemical Processes				
Kraft	—	Sodium hydroxide + sodium sulfide	None	45–55
Sulfite	—	Calcium bisulfite in sulfurous acid[c,d]	None	40–50
Magnefite	—	Magnesium bisulfite in sulfurous acid[e]	None	45–55
Soda	—	Sodium hydroxide	None	40–50
Soda-oxygen	—	Sodium hydroxide + oxygen	None	45–55
Soda-anthraquinone	SAq	Sodium hydroxide + anthraquinone	None	45–55
Dissolving Processes				
Prehydrolysis kraft	—	Steaming and kraft (two-step process)	None	35
Acid sulfite	—	Acid sulfite (Ca, Na)	None	35

Source: Data taken from references 1 through 5. [a]Sodium sulfite or sodium hydroxide, 2 to 7% of wood. [b]Also chemical treatment after fiberizing. [c]Also sodium, magnesium, ammonia. [d]pH 2. [e]pH 5.

paper. Therefore, for all but low-grade pulps, bark should be removed as well as possible, and this removal must be very thorough in the case of groundwood and sulfite pulps if the finished paper is to appear clean.

Debarking usually is done in a drum barker, where the bark is removed by the rubbing action of logs against each other in a large rotating drum. Hydraulic barkers using high-pressure water jets are excellent for large logs and are common on the West Coast. Mechanical knife barkers are becoming more common and are used extensively in smaller operations because of their relatively low capital cost. Also they have found widespread use in sawmills to debark logs prior to sawing so that the wood wastes can be used to produce pulp.

Wood cut in the spring of the year during the active growing season is very easy to peel. Much of the spruce and fir cut in the North is still hand-peeled during this season and usually represents the optimum in bark removal.

Fig. 7.11. Relative wood pulp production by grade.

The standard log length used in the Northeast is 48 inches, whereas 63 inches is common in the South. Wood generally is measured by log volume, a standard cord being considered to contain 128 cubic feet. Large timber on the West Coast generally is measured in board feet of solid volume. Measurement and purchase of wood on a weight basis are practiced, and are desirable because weight is directly related to fiber content.[1]

The growth in the use of wood residues has been substantial. By barking the sawlogs, the slabs, edgings, and other trimmings that formerly were burned now can be used to make pulp. Almost 28 percent of the wood used by the pulp industry is classified as waste wood. Several mills have been built that use no logs whatsoever but depend on residuals from satellite sawmill operations. Special sawmilling equipment has been developed to produce sawdust of a proper size so that it too can be used.

Wood used in producing groundwood or mechanical pulp requires no further preparation after debarking, but that used in the other chemical processes first must be chipped into small pieces averaging one-half to one inch in length and about one-eighth to one-quarter inch in thickness. Chipping is accomplished with a machine consisting of a rotating disc with knives mounted radially in slots in the face of the disc. Modern chippers have up to 12 knives; the ends of the logs are fed against the disc at about a 45° angle.

Mechanical Pulping

Mechanical pulping, as the name implies, involves mainly grinding and abrading action on wood to produce a mechanical pulp—the primary component of newsprint. Modern mechanical pulping includes stone groundwood pulping (SGW), in which bolts of wood are pressed against a revolving grindstone, and refiner mechanical pulping (RMP), in which chips are passed between single- or double-rotating plates of a vertical-disc attrition mill. Recent developments in stone grinding include application of pressure to the grinder (PGW) and control of temperature.

At one time, natural sandstone was used for the grindstone, but modern stones are either silicon carbide or aluminum oxide grits in a vitrified clay binder. Thus the characteristics of the stone can be varied to produce pulps "tailor-made" to fit their desired end use. Groundwood pulps for roofing or flooring felts must be extremely coarse and free-draining, so they require a stone with large grits, whereas newsprint pulps are very fine and require the use of small grits.

Pulp characteristics also can be varied by changing the stone surface pattern, the stone

speed, the pressure of the logs against the stone, and the temperature of the ground-pulp slurry. Generally a coarser and more freely draining pulp is obtained with a coarse surface pattern and high speed, pressure, and temperature. The type and the condition of the wood also are factors, but groundwood pulps usually are made from the coniferous or long-fibered species because the deciduous or short-fibered species give very weak pulps.

Basic changes in mechanical pulping technology include pretreating chips with chemicals, steam, or both. These developments started 30 years ago when chips were pretreated with caustic soda, in a process called chemimechanical pulping (CMP). Presteaming and pressure refining of chips gives a thermomechanical pulp (TMP). The chips are steamed at 40 psi for 2 to 4 minutes and sent to a refiner modified to operate under pressurized conditions. When chemical pretreatment and pressure steaming are combined, the pulp is referred to as chemithermomechanical pulp (CTMP). Sulfite solutions commonly are used to produce CTMP, but there are many variations of these processes. These treatments are employed to improve pulp quality. The steam and chemicals aid fiberizing by giving a less-damaged fiber, which makes the final paper stronger. The pulps are sufficiently improved over stone groundwood that they can be used to partially substitute for chemical pulps. Newsprint consisting of 100 percent TMP or CTMP is produced commercially.

Groundwood or mechanical pulp is low in strength compared to the chemical pulps. It is composed of a mixture of individual fibers, broken fibers, fines, and bundles of fibers. Papers made from groundwood also lose strength and turn yellow with time. Thus, groundwood pulps are used only in relatively impermanent papers such as for newsprint, catalogs, magazines, and paperboard. Groundwood papers have excellent printing qualities because of high bulk, smoothness, resiliency, and good ink absorption. Newsprint contains about 80 percent groundwood, and the other publication paper grades contain about 30 to 70 percent, with the remainder a chemical pulp to provide greater strength. Groundwood is the cheapest pulp made and also utilizes the entire wood, giving close to a 100 percent yield.

Chemical Pulping

The objective of chemical pulping is to solubilize and remove the lignin portion of wood so that very little mechanical treatment is necessary to fiberize the wood. Many variations are used throughout the world, the simplest classification of pulping methods depending on whether they are acidic or alkaline. Each has its own specific advantages and disadvantages, but the kraft (alkaline) process accounts for over 90 percent of all chemical pulp produced in the United States. All present commercial processes use aqueous systems under heat and pressure although several organosolv systems are at the demonstration-plant stage.

The lignin-removal and fiber-separation mechanisms involved in chemical pulping are alkaline hydrolysis (lignin bond cleavage) and formation of soluble sulfur-containing lignin (thiolignins) in kraft pulping. The chemical mechanisms in sulfite pulping also involve lignin bond hydrolysis and formation of soluble sulfur-containing lignin derivatives, termed lignosulfonates.[5] Polysaccharide reactions are secondary to fiber separation in chemical pulping although polysaccharide retention, solution, or modification profoundly affects pulp yield and properties. Neutral sodium sulfite reactions in neutral sulfite semichemical (NSSC) pulping also are based on sulfonation for partial dissolution of lignin and a partial modification of the fiber bond to permit a clean separation of the fibers by mechanical action in disk refiner mills. The pulping chemicals are applied in water solutions (white liquor) at various concentrations and ratios of liquids to solids under different conditions of temperature and pressure to achieve the desired degree of delignification.

Kraft Process. The kraft (sulfate) process uses a mixture of sodium hydroxide and sodium sulfide as the active chemical. Although

TABLE 7.3 Typical Kraft Pulping Conditions

Pressure	100–110 psig
Temperature	170–175°C
Time	2–3 hrs
Alkali charge	15–25% of weight of wood (calculated as Na_2O but consisting of approximately $5NaOH + 2Na_2S$)

Liquor to wood ratio is 4 to 1 (by weight).

the designation sulfate process is misleading, it sometimes is used because sodium sulfate is the makeup chemical in the process. The word "kraft" now is mainly used to describe this process, derived from the Swedish or German word for strength, as this process produces the strongest pulp. Historically, sodium hydroxide alone (soda process) first was used as the alkaline pulping agent, but all these mills have converted to the kraft process because the other pulp is weak and inferior to kraft pulp.

The pulping (cooking) process traditionally was performed on a batch basis in a large pressure vessel called a digester. Conditions vary depending upon the type of wood being pulped and the quality of end product desired. Typical conditions for kraft cooking are listed in Table 7.3.

Digesters are cylindrical in shape with a dome at the top and a cone at the bottom. Ranging in size up to 40 feet high and 20 feet in diameter, the largest will hold about 7000 cubic feet of wood chips (about 35 tons) for each charge. The chips are admitted through a large valve at the top, and at the end of the cook they are blown from the bottom through a valve to a large blow tank. During the cook the liquor is heated by circulation through a steam heat exchanger, which also avoids the dilution of the cooling liquor that would occur from heating by direct injection of steam.

The development of the continuous digester (Fig. 7.12) significantly facilitated the use of kraft pulping. Chips are admitted continuously at the top through a special high-pressure feeder, and the cooked pulp is withdrawn continuously from the bottom through a special blow unit. Recent installations range in size up to 150 feet high and are capable of

Fig. 7.12. Typical Kaymr continuous cooking system for kraft pulp. (*Courtesy Kamyr, Inc.*)

producing about 1000 tons of pulp per day in one unit. Cooking liquors and conditions are approximately the same as for the batch digesters. These units offer both good economics in the production of pulp and a quality advantage compared to the batch digester. However, because the capital investment for these units is somewhat higher, both systems are installed for new mills.

Because of the high alkali charge, the chemicals must be recovered and reused. This process also alleviates pollution problems because the yield of pulp is only about 45 percent of the original wood weight, and the organic residues must be eliminated. After being cooked in the digester, the pulp is washed in a countercurrent rotary vacuum washer system using three or four stages. The pulp then is ready for bleaching or for use in papers such as grocery bags where the brown color is not objectionable.

The separated liquor, which is very dark, is known as "black liquor." It is concentrated in multi-effect evaporators to 60 to 65 percent solids. At this concentration the quantity of dissolved organic compounds from the wood (lignin and carbohydrate degradation products) is sufficient to allow the liquor to be burned in the recovery furnace.

By controlling the amount of excess air admitted to the furnace and the temperatures, the organics in the liquor can be burned. The inorganics collect on the bottom of the furnace as a molten smelt of Na_2CO_3 and Na_2S. Sodium sulfate is added to the liquor as makeup and is reduced to Na_2S by carbon. After dissolving in water, this mixture (called "green liquor") is reacted with slaked lime:

$$Na_2CO_3 + Ca(OH)_2 \rightarrow 2NaOH + CaCO_3$$

The Na_2S does not react with the lime, so the resultant mixture of NaOH and Na_2S (called "white liquor") can be reused to pulp more wood. The $CaCO_3$ sludge is filtered off, burned in a lime kiln, and reused. Thus, the chemical system is a closed one, as shown in Fig. 7.13, and this minimizes costs and pollution.

The kraft process has had a serious problem with air pollution due to the production of

Fig. 7.13. Diagram showing cyclic nature of kraft recovery process. (*From* Pulp and Paper, *3rd ed., Casey, J. P., Ed. Wiley-Interscience Pub., New York. Copyright © 1980 by John Wiley & Sons, Inc. and reproduced by permission of the copyright owner.)*

hydrogen sulfide, mercaptans, and other vile-smelling sulfur compounds. In recent years, the use of various techniques such as black-liquor oxidation, improved evaporators and furnaces, and control of emissions has greatly improved this situation. However, older mills are being forced to expend large amounts of money to bring their operations up to the environmental standards.

Practically any kind of wood can be pulped by the kraft process; and because it produces the strongest pulps with good economies, it has grown to be the dominant process in the world. Although the odor problem is very severe, it does not appear that this process will be supplanted in the near future; instead, it will be improved and modified. When pulping resin-rich woods such as southern pine, the kraft process yields turpentine and tall oil as valuable by-products. The steam generated in the recovery furnace is almost enough to make the pulp mill energy self-sufficient.

A recent discovery that small amounts of anthraquinone (0.05–0.25% on wood) added to alkali liquors can enhance delignification and produce kraft quality pulp has been of great interest. Although the price of anthraquinone (AQ) is high and the chemical is not recovered, only small quantities are required, and the benefits outweigh the costs sufficiently that a number of mills around the world now routinely add AQ to the kraft pulpwood digester.

Sulfite Process. The sulfite process uses a cooking liquor of sulfurous acid and a salt of the acid. Although calcium was the most widely used base at one time, it has been supplanted by sodium, magnesium, and ammonia. Lignin will react with the bisulfite ion (HSO_3^-) under acidic conditions to form lignosulfonates that are soluble in water. For many years this was the preferred process because it produced pulps of light color that could easily be bleached, it used cheap chemicals in fairly limited amounts so that no recovery was necessary, and it was a relatively simple process to operate.

Although the production of sulfite pulp has remained relatively constant for the last 30 years, the rapid growth of kraft pulping has reduced sulfite's share to less than 10 percent of the chemical pulp produced. There are several reasons for this, but the primary ones are the inability to cook resinous woods such as pine, problems in producing strong pulps from hardwoods, and of the greatest importance, the lack of a cheap and simple recovery system to reduce water pollution problems. However, this process produces pulps with special qualities such as high alpha-cellulose grades for rayon, so it will continue to be used.

Initially calcium was the preferred base because it was cheap and convenient to use. However, no recovery system is available for this base, so most calcium-base mills either have ceased operation or have converted to sodium, magnesium, or ammonia, for which recovery systems are available.

Regardless of the base used, the initial step is the burning of sulfur to produce sulfur dioxide (SO_2). The air supply to the burner must be carefully controlled, as too much air will enhance the formation of sulfur trioxide (SO_3) and subsequent production of sulfuric acid (H_2SO_4), which is very undesirable. The gas also must be cooled quickly from 1000°C, leaving the burner at below 400°C, to minimize formation of SO_3. After cooling to 20 to 30°C, the SO_2 gas must be absorbed in water and reacted with the proper base to form the cooking liquor.

For calcium-base liquor, the gas is passed through towers packed with limestone with water flowing down through the tower. Because of the limited solubility of calcium bisulfite [$Ca(HSO_3)_2$], the pH of the liquor is very low (about 2), and free sulfurous acid is present. This usually is called the acid sulfite process. As mentioned before, calcium-base mills have essentially disappeared in the United States.

The so-called soluble bases now are used, with each having certain advantages. Because solutions of sodium, magnesium, and ammonium bisulfite are all soluble at pH 4.5, the current practice is to pulp at the higher pH, which is usually called bisulfite pulping. Extremely long cooking times (7–10 hours)

are necessary with acid sulfite, whereas 4 to 5 hours is sufficient with bisulfite.

Sodium base is the easiest to prepare (Na_2CO_3 or NaOH usually is used as the makeup chemical) and gives the highest-quality pulp; however, recovery processes, though available, are complicated and expensive. Magnesium base [from $Mg(OH)_2$] is somewhat more difficult to handle, but recovery systems are available, and the majority of the sulfite pulp now is produced from this base. Ammonium base (from NH_4OH) was used in the past; the ammonia cannot be recovered, but the liquor can be evaporated and burned without leaving any solid residue, thus reducing water pollution. As long as aqueous ammonia remains low in price, this process will be attractive because the SO_2 can be recovered from the waste gases by passing them through a wet scrubber flooded with fresh ammonium hydroxide. The various sulfite-based processes are summarized in Fig. 7.14.

Batch digesters usually are used in the sulfite process. Cooking temperatures are lower (140–150°C), and times are longer than in the kraft process, with pulp yields about the same as in the kraft process. Spruce and fir are the preferred species for cooking by the sulfite process because they produce relatively strong, light-colored pulps. About 20 percent of newsprint consists of this type of pulp that has not been bleached. Thus the sulfite industry is concentrated in Canada, the northern United States, and the Pacific Coast, where the supplies of spruce and fir are greatest and the largest quantities of newsprint are produced.

A large amount of research has been done on developing products from the waste sulfite liquor, and some success has been achieved. Vanillin, alcohol, and torula yeast can be

	Acid Sulphite	Bisulfite	Neutral Sulfite Semichemical
pH	1.5	4 - 5	8 - 9
Chemicals	H_2SO_3, $NaHSO_3$	$NaHSO_3$	Na_2SO_3, Na_2CO_3
Base	Ca, Na, Mg, NH_4	Na, Mg, NH_4	Na, NH_4
Rate of Cooking (Given temp.)	← Decreasing →		
Temperature	140°C	170°C	170 - 180°C
Brightness	← Slight increase →		← Decrease →
Strength Characteristics (Given yield)	← Increasing →		
Species	Nonresinous	Hardwood & Softwood	Hardwood
Equipment Corrosion	← Decreasing →		

Fig. 7.14. Summary of sulfite pulping processes.

produced as by-products, and the lignosulfonates are used as viscosity modifiers in drilling muds and for similar purposes. However, the majority of waste liquor is burned to recover the cooking chemicals and the heat values.

Semichemical Pulping. Various combinations of chemical and mechanical treatments have been used to produce pulps with specific properties. Mild chemical treatments to give partial delignification and softening are followed by mechanical means to complete fiber separation.

The neutral sulfite semichemical (NSSC) process is one in which wood chips, usually from hardwoods, are cooked with Na_2SO_3 liquor buffered with either $NaHCO_3$, Na_2CO_3, or NaOH to maintain a slightly alkaline pH during the cook. Unbleached pulp from hardwoods cooked to a yield of about 75 percent is widely used for the corrugating medium. Although bleachable pulps can be produced by this process, they require large quantities of bleaching chemicals, and the waste liquors are difficult to recover. Currently many NSSC mills are located adjacent to kraft mills, and the liquors can be treated in the same furnace. Thus the waste liquor from the NSSC mill becomes the makeup chemical for the kraft mill, solving the waste problem. NSSC hardwood pulp is the premier pulp for the corrugating medium and cannot be matched by any other process.

Chemimechanical pulps usually are produced by soaking the chips in solutions of NaOH or Na_2SO_3, and then refining in disc refiners to produce a groundwood-type pulp. Chemical consumption is very low, and yields are usually 85 to 95 percent.

Solvent Pulping. The concept of using solvents to dissolve lignin was first explored by Peter Klason in 1893. In the 1930s, Kleinert developed and patented the first solvent (organosolv) pulping process based on 50 percent aqueous ethanol.[7] Until recently there was little incentive to pursue alternatives to the kraft process because energy and chemicals were reasonably cheap, there were few environmental regulations, and there was little interest in by-products. However, new, strict environmental regulations, rising energy and chemical costs, and the very large capital investments needed to construct new kraft mills have stimulated research and investment in alternate solvent pulping processes.

With organosolv pulping, wood chips are contacted with an aqueous solvent mixture and cooked (digested) for a period of time at elevated temperatures. Digestion pressures range from atmospheric to 500 psig. Cooking initially releases acetic and formic acids from ester functionalities in the wood, and these acids further promote hydrolysis of lignin and hemicelluloses to low molecular weight species. Catalysts such as mineral acids (HCl), organic acids (acetic, formic), sulfonic acids, and Lewis acids and bases ($AlCl_3$, $Fe_2(SiO_4)_3$, $Mg(SO_4)$, $CaCl_2$) can be employed to promote delignification. Cellulose pulp is filtered from the liquor, and the solvent is recovered, usually by distillation. Lignin becomes insoluble at this point and precipitates from the liquor, leaving an aqueous solution of primarily hemicellulosic sugars. Conditions for a variety of organosolv processes are compared in Table 7.4.[8]

Two processes that have reached the demonstration-plant scale are the Alcell and the MD-Organosolv processes, both of which are basic modifications on the original process patented by Kleinert.[7] In the Alcell process preheated chips (80°C) are packed in the digester using steam, and the steaming is continued to displace air from the chips. Preheated solvent that has been previously used as a wash liquor for two earlier batches is pumped in and rapidly brought up to pulping temperatures of 190 to 200°C, corresponding to an operational pressure of 500 psi. The liquor is continuously circulated in the digester during the pulping period. At the end of this period, the liquor in the digester is displaced by wash liquor that was used in the second washing stage of the previous batch. The displaced liquor from the digester flows to the lignin and sugar recovery system. The third-stage wash liquor is displaced by fresh liquor and flows into an intermediate

TABLE 7.4 Solvent Pulping Process Characteristics

Process Name Developer	*Solvent Used*	*Additional Reactants*	*Digestion Temperature*	*Digestion Time*	*Reference*
Acetic Acid (Herdle, Young)	90% acetic acid[a]	0.2% H_2SO_4	150°C	4 hr	(9)
Acetosolv (Nimz)	95% acetic acid	0.1% HCl	110°C	3–5 hr	
Ester (Young)	50–75% acetic acid[a]	Ethyl acetate[d]	170°C	1 hr	(10)
Formic Acid (Jordan)	80% formic acid	Unspecified organic catalyst	105°C	45 min	(11)
Peroxoic Acid (Poppins, Sundquist)	80–100% formic acid and hydrogen peroxide	None	70–80°C	60 min	
Hydrotropic Pulping (Springer)	2.0 M xylene-sulfonic acid[a]	None	100°C	60 min	(12)
Battelle-Geneva Process (Sachetto)	Aqueous phenol[b]	Acid catalyst	100°C	30–60 min	(13)
Sulfolane (Clemont)	50% sulfolane	None	175°C	2 hr	(14)
Acetone (DeHaas and Lang)	Aqueous acetone[b]	NH_3/Na_2S	160–190°C	1–2 hr	(15)
Glycol (Burkart, Quigley)	20–100% glycol	Mineral, salicylic, sulfonic acids	100–205°C	1–3 hr	
Phenol (Sweers)	20–50% phenol	HCl. sulfonic acids	80–205°C	1–4 hr	
Cresol (Sakakibara)	20–80% cresol	Cuprous oxide, acetic acid, AQ	160–190°C	1–4 hr	
Amine (Wallis)	25–90% mono- or triethanol amine	None, AQ	165–200°C	1–2 hr	
Kleinert Process	20–75% ethanol[a,c] usually 50%	Uncatalyzed or metal salt catalysts	160–195°C	30–60 min	(7, 16, 17)
Alcell (Pye)	50% aq. ethanol	Uncatalyzed	195°C	1–2 hr	
AEM (Paszner)	Aqueous methanol[b]	Alkali earth salt catalysts	195–210°C	25–45 min	(18)
MD-Organosolv (Edel)	Aqueous methanol	NaOH	170–195°C	45 min	(19)
ASAM (Patt)	30–50% aq. methanol	NaOH, Na_2SO_3, and AQ (all)	180°C	2–3 hr	

[a]Other compound is water. [b]Concentration not reported. [c]*n*-Butanol also has been employed. [d]A three-component (ethyl acetate/acetic acid/water) pulping process.

storage tank for use in the second stage. The second-stage wash liquor is drained into another storage tank for use in the digester.

At this point, the digester contains soaked, delignified chips and alcohol–water vapor at pulping temperature and pressure. It is now depressurized, and the departing vapor is condensed for reuse as fresh pulping liquor. Finally, the alcohol remaining in the digester is driven off by steaming. The pulp is diluted with water and pumped out of the digester for cleaning and bleaching treatments. The first-stage spent liquor entering the by-product recovery area is first flashed and then diluted with process water to precipitate the dissolved lignin. After settling, the solid lignin is separated by centrifugation, washed, and dried. The filtrate enters a distillation tower, where alcohols and some acetic acid and furfural are recovered. Finally, the remaining aqueous liquor is further evaporated to a sugar syrup.[20]

Currently a demonstration plant is producing 15 metric tons of organosolv pulp a day in New Brunswick, Ontario. A comparison of the strength properties of Alcell and kraft pulps is given in Table 7.5. The process as described above is only suitable for hardwoods.[21]

The principles of the MD-Organosolv process, developed in Germany, are illustrated in Fig. 7.15. The process consists of first impregnating softwood chips with 50 percent aqueous methanol before they enter the

TABLE 7.5 Comparison of Properties of Allcell and Kraft Hardwood Pulps*

Property	*Alcell Pulp*	*Commercial Kraft Pulp*
Tensile (km)	7.47	7.40
Tear ($mN \cdot m^2/g$)	7.20	6.75
Burst ($kPa \cdot m^2/g$)	5.08	5.18
Brightness (ISO)	88.7	89.6

*Pulps 50% maple, 35% birch, 15% poplar; beaten to 400 Canadian Standard Freeness.
Source: McCready, M. (Ed.), *Paper Age* (Mar. 1991).

first-stage continuous reactor. The first pulping stage is operated under 40-bar pressure at 195°C and removes approximately 20 percent of the lignin and a major part of the hemicelluloses from the chips. The chips then are transported hydraulically to the second-stage digester where alkali is charged at 15 to 20 percent based on wood, and the methanol concentration is reduced to about 30 percent. The second stage is operated at 170°C, and the combined residence time in both reactors is only 45 minutes. The strength properties of the pulps produced are only slightly lower than those of corresponding kraft pulps. The recovery of organics and NaOH from the second-stage liquor is performed in a novel manner. After methanol evaporation, the liquor enters an electrolytic cell that not only separates NaOH but also precipitates lignin in a filterable form. The 5 ton/day demonstration plant will be scaled up to an 85 ton/day installation to supplement kraft pulp demands.[20]

Screening and Cleaning of Wood Pulp

The desired pulp fibers are usually between 1 and 3 mm in length with a diameter about one-hundredth as large. Any bundles of fibers or other impurities would show up as defects in the finished paper and must be screened out. Wood knots usually are difficult to pulp and must be removed.

Screening is usually a two-stage process with the coarse material being removed by screens with relatively large perforations (1/4 to 3/8 in.). Additional fine screening is done with screens using very small (0.008 to 0.014 in.) slots to ensure removal of oversized impurities. Screen size openings will depend on the species of wood being processed and the desired quality of the end product. Because of the tendency of the fibers to agglomerate when suspended in water, it is customary to screen at very low solids (consistencies) of about 0.5 percent fiber and 99.5 percent water.

To meet the ever increasing demands for cleaner pulps, centrifugal cyclone cleaners have come into almost universal use. The screened pulp is pumped through these units at low consistencies and high velocities. The fiber slurry enters the cone tangentially at the top, and a free vortex is formed with the velocity of the flow greatly increased as the diameter of the conical section is reduced. Heavier particles of sand, scale, or other dirt are forced to the outside of the cleaner and are discharged from the bottom tip through a small orifice. Because of the velocity gradients existing in the cone, the longer fibers (75–95%) are carried into the ascending center column and are discharged through the larger accept nozzle at the top. In a properly designed and operated unit, a shape separation also is made so that round particles, even though of the same specific gravity as the good fibers, will be discharged as rejects through the bottom orifice. In this way, small pieces of bark also are removed. To reduce the quantity of rejects to an acceptable level, they in turn are processed through a second, a third, or even a fourth stage of cleaners, thus holding the final loss of pulp to about 1/4 to 1 percent of the feed, depending upon quality demands and the dirt level of the incoming pulp.

Bleaching of Wood Pulp

The color of unbleached pulp ranges from cream or tan for the sulfite process to dark brown for the kraft process. Although about 75 to 90 percent of the lignin has been removed by the pulping process, the remainder, along with other colored degradation products, must be removed by bleaching.

Although it is possible to improve the brightness (whiteness) of the pulp in one stage,

Fig. 7.15. lowsheet of MD-Organosolv pulping system. (*From Aziz, A., and K. V. Sarkanen,* Tappi **72**, *No. 3, 169–175 (1989); with permission of Tappi, Atlanta, GA.*)

the economical achievement of high brightness requires the use of several stages. Current practice uses combinations of chlorination with elemental chlorine (C), alkaline extraction with sodium hydroxide (E), and various oxidative stages using sodium or calcium hypochlorite (H), chlorine dioxide (D), or hydrogen peroxide (P). The pulp is washed between each stage to remove solubilized impurities. Many combinations are possible, and each mill selects the sequence that fits its requirements the best. The sequences most commonly utilized are:

CEH	CED
CEHD	CEDED
CEHDP	CEHDED

TABLE 7.6 Typical Conditions for Bleaching Kraft Pulp

Stage	*% Chemical on Pulp*	*Time (min)*	*Temperature (°F)*
Chlorination	5–6	30	70
Extraction	$2\frac{1}{2}$–3	60	140
Hypochlorite	1–2	90	105
Chlorine dioxide	$\frac{1}{2}$–1	240	160

The greater the number of stages, the higher the quality of the final pulp but at increased cost. Sulfite pulps are much easier to bleach and usually use only the three- or four-stage sequences, whereas kraft pulps require additional stages. Typical conditions for bleaching a kraft pulp are given in Table 7.6.

Chlorine and caustic are purchased, but chlorine dioxide must be generated at the site by using sodium chlorate as the basic chemical. Effluents from bleach plants are a source of great concern in regard to pollution. Large quantities of water are discharged (typically 10,000–20,000 gal/ton of pulp) that are high in color, especially the water coming from the chlorination and extraction stages.

New developments involve the replacement of the chlorination and extraction stages with a single stage involving gaseous oxygen and sodium hydroxide. The pulp is dewatered to about 15 to 25 percent solids, treated with about 4 to 6 percent NaOH, and passed through a reactor in a fluffed condition using oxygen gas at about 150 psi. The pulp consumes about 1 to 2 percent oxygen in about 15 minutes and leaves in a semibleached condition. After washing, the use of conventional chlorine dioxide stages will produce the desired level of whiteness. The advantage of this process is that the effluent contains no chlorides and can be used as process water in the pulp mill. Any sodium and organics in the effluent then will enter the kraft recovery cycle and serve as makeup, thus reducing the pollution load.

These bleaching sequences are designed to remove lignin, yielding a highly purified fiber consisting only of carbohydrate material. For producing high-yield pulps such as groundwood, where all of the lignin is retained in the pulp, it is not possible to use these systems. However, extremely high brightness is not required in this case; thus, some improvement is attained by using one stage with either peroxide or hydrosulfite (dithionite). No yield loss is encountered as the action of both of these is merely to decolorize the pulp rather than to remove any impurities. Usually about 0.5 percent of either of these chemicals will give a noticeable increase in brightness, and they are widely used to upgrade the quality of groundwood.

Biotechnology—Biopulping and Biobleaching

The pulping of wood is at present based on either mechanical or chemical methods or combinations thereof, as previously described in this chapter. The interfiber lignin bond is broken down by the mechanical and chemical treatments to free the cellulose fibers for papermaking. In the forest, white rot fungi perform a similar task on wood left behind. The enzymes of the fungi do the work of lignin degradation.

It should be possible to isolate these specific enzymes from the fungus for use in biological pulping or biopulping. Indeed, researchers at several laboratories are attempting to isolate lignin-degrading enzymes (lignases) from the white rot fungus *Phanerochaete chrysoporium.* They are quick to point out, however, that there probably is a lignase complex, in which several enzymes work together to degrade lignin. Once specific enzymes are isolated, work can commence on producing the enzymes in greater quantities through genetic engineering. Another drawback is the slow rate of reaction.

The initial applications of lignases would be to degrade the lignin partially before mechanical treatment so that the process would require less energy, and the pulp would have more strength after this biomechanical pulping. Biobleaching also would be possible for brightening or whitening pulp fibers, in lieu of the toxic chlorine compounds utilized at present by the industry. A particularly exciting

application would be to use these types of enzymes for removal of lignin pollutants from waste effluents. Biotechnology should lead to safer and cleaner methods for pulping and bleaching.

Recycling

Worldwide, approximately 32 percent of wastepaper (73 million tons) is recycled. However, the rate of recycling varies considerably between countries, as shown in Fig. 7.16. The United States is a major exporter of wastepaper—over 4 million metric tons in 1990. The majority of the wastepaper exported goes to "fiber-poor" countries that have much less virgin fiber than does the United States and therefore recycle much greater quantities of paper.[5]

A variety of problems are associated with paper recycling, such as collection, distribution, and wild cyclic swings in the market. However, with landfill sites now at a premium and paper representing about 40 percent of municipal solid waste, it makes good sense in the long run to promote paper recycling, which reduces landfill needs and the consumption of virgin timber. The U.S. paper industry has set the goal of a 40 percent wastepaper collection rate by the year 1995. This means that large volumes of wastepaper will be available for reuse in the future.

There are a wide range of different grades of wastepaper available, depending on the source and the extent of separation. As the name implies, wastepaper designated as "direct pulp substitutes" is utilized with little treatment before reslushing in a hydrapulper. The direct pulp substitutes are the highest grade of wastepaper.

Fig. 7.16. Comparison of paper recycling in selected countries.

The majority of recycled paper (about 75%) is used with no attempt to remove inks, dyes, or pigments from the paper. The resultant pulp is of rather poor color and quality, and is used primarily as filler stock in paperboard. The bulk grades are the largest-volume wastepaper source.

Deinked grades of wastepaper require special techniques and equipment to remove inks, coatings, adhesives, and so on. The deinking process is complicated and time-consuming; depending on the quality of deinked pulp required, the process may involve a number of combinations of washing, flotation, dispersion, screening, and use of cyclone cleaners. The newer noncontact inks present special removal problems, as do wet-strength agents, adhesives, and the so-called stickies. Stickies are made from hot-melt adhesives (vinyl acetate polymers and copolymers), pressure-sensitive adhesives (styrene-butadiene), and lattices (natural and synthetic rubber). Tackifiers and waxes also are usually included in these adhesive formulations. A variety of additives are used to help remove stickies and other contaminants from wastepaper, including solvents, nonionic and cationic surfactants, zirconium compounds (to reduce tackiness), and talc. Of course, all of the additional steps and additives add to the expense of recycling. There is also the problem of waste disposal from the deinking process, which must be properly handled. Table 7.7 shows the projected consumption of wastepaper according to paper grade for 1995.[22]

Stock Preparation

Stock preparation in a paper mill includes all intermediate operations between preparation of the pulp and the final papermaking process. It can be subdivided into (1) preparation of the "furnish" and (2) "beating" or "refining." Furnish is the name for the water slurry of

TABLE 7.7 Projected Consumption of Wastepaper in Paper and Paperboard Manufacture for 1995

	Total Production	*Wastepaper (%)*
Production	93,250	29.9
Paper	47.450	17.1
Newsprint	7,400	36.5
Printing and writing	28,800	7.0
Packaging	4,850	5.2
Tissue	6,400	49.4
Paperboard	44,200	43.0
Unbleached kraft	22,000	20.0
Bleached kraft	48,000	0
Semichemical	6,400	40.6
Recycled	11,000	109.1
Construction paper and board	1,600	50.0

fiber and other chemicals that goes to the paper machine. Beating or refining refers to the mechanical treatment given to the furnish to develop the strength properties of the pulp and impart the proper characteristics to the finished paper.

Cellulosic fibers are unique in that, when suspended in water, they will bond to each other very strongly as the water is removed by filtration and drying, with no need for an additional adhesive. This is so because of the large number of hydrogen bonds which form between the surfaces of fibers that are in close contact as the water is removed. This bonding is reversible and accounts for the well-known fact that paper loses most of its strength when wet. If paper is suspended in water and agitated, it will separate into the individual fibers, a behavior that allows the easy reuse of wastepaper or the processing-waste from the paper mill.

In order to enhance the bonding capability of the fibers, it is necessary to mechanically beat or refine them in equipment such as beaters, jordans, or disc refiners. This treatment of the pulp slurry at about 3 to 6 percent consistency is done by passing the pulp between the two rotating surfaces of the refiner. These surfaces contain metal bars and operate at very close clearances. As the fibers pass between the bars, they are made more flexible, and a larger surface for bonding is developed by the mechanical action.

This refining brings about fundamental changes in the pulp fibers and increases the degree of interfiber bonding in the final sheet of paper. Thus the final properties of the paper can be significantly changed by varying the degree and type of refining. As additional refining is performed, properties such as tensile strength, fold, and density are increased while tear resistance, opacity, thickness, and dimensional stability are decreased. Thus, the proper refining conditions must be selected to bring out the desired properties without detracting too much from other properties.

The furnish of a paper machine varies widely, depending on the grade of paper being made. Newsprint usually consists of about 80 percent groundwood and 20 percent chemical fiber (sulfite or semibleached kraft). Bag papers and linerboard are usually 100 percent unbleached softwood kraft. Printing papers are made from bleached pulps and contain both hardwoods and softwoods. By selecting the proper pulps and refining conditions, a wide variety of paper qualities can be achieved.

The paper industry is a large user of chemicals because relatively small quantities of additives can materially change the properties of paper. The use of 1 to 2 percent rosin size and 2 to 3 percent alum [$Al_2(SO_4)_3$] will greatly increase the resistance of paper to penetration by water or ink. Pigments such as kaolin clay, calcium carbonate, and titanium dioxide are added in amounts up to 15 percent to increase opacity and give a better printing surface. Organic dyes and colored pigments are added to produce the highly colored papers used for business and printing papers. Other additives such as wet-strength resins, retention aids, and starch can be used to give particular properties that are needed. Thus, in order to produce the wide variety of grades of paper now available, the papermaker selects the proper pulps, refining conditions, and additives, and then combines the pulp and additives before sending them on to the paper machine for the final step in the process.

Papermaking Process

Some paper mills are not integrated with pulp mills, and it is necessary for these mills to use dried, baled pulp manufactured at a separate location. Many mills making limited quantities of highly specialized papers fall into this category, as it allows maximum flexibility in selecting the optimum pulps for a particular paper grade. However, the papermaking process is the same regardless of the source of pulp.

After the furnish has been prepared with the proper refining treatment and additives, it is stored in the machine chest and then fed continuously into the paper machine system. A refiner or jordan is placed in this line to give the paper machine operators the opportunity to make small adjustments in the quality of the furnish as needed to give the desired paper properties. Screens and centrifugal cleaners also are included to ensure a high-quality paper.

The papermaking process is essentially a system whereby the pulp is diluted to a very low consistency (about 0.5%) and continuously formed into a sheet of paper at high speeds, and then the water is removed by filtration, pressing, and drying. The basic units of the fourdrinier paper machine are diagrammed in Fig. 7.17, a picture is shown in Fig. 7.18.

The section of the paper machine where the paper is formed is referred to as the "wet end." The fourdrinier machine is characterized by a headbox that allows the diluted stock to flow through an orifice (slice) onto the flat moving wire. This is actually an endless wire belt that returns on the underside of the machine, thus allowing the process to proceed continuously. Because a low headbox consistency (about 0.5%) is necessary for good formation, the volumes of water handled are very large (about 10,000 gal/min for a machine producing 300 tons/day). Much of the water is removed through the wire by the action of the table rolls and foils that support the wires in the forming area immediately following the headbox. At this point the stock consistency has been raised to about 2 percent,

Fig. 7.17. Diagram of a fourdrinier paper machine. (*Courtesy Hammermill Paper Co.*)

Fig. 7.18. Modern high-speed paper machine. (*Courtesy Beloit Corp.*)

and the stock no longer drains freely. By passing it over suction boxes operating at fairly high vacuum (6–8 in. Hg), the consistency is raised to about 15 to 20 percent. A dandy roll (covered with woven wire) rotating on top of the wet paper is used to improve the formation and can impact a watermark if it contains the proper pattern. The suction roll after the suction boxes has several functions: it removes additional water; it serves as the driving roll for the wire, and it serves as the point at which the paper separates from the wire and passes into the press section while the wire returns to pick up additional pulp at the headbox.

Modern fourdrinier machines are available up to a width of 360 inches and can operate at speeds up to 3000 ft/min. Newsprint machines are usually the widest and fastest, whereas those making heavier grades such as linerboard run somewhat more slowly. In the United States, there are several machines capable of producing over 1000 tons/day of linerboard. More specialized grades such as bond and printing papers are usually produced at a lower speed on narrower machines, and 150 to 250 tons/day would be considered a high output. Many specialized grades such as filter paper and tracing paper are produced on very small, slow machines producing only a few tons per day. Machines making tissue paper for sanitary purposes use modifications of the standard fourdrinier to produce tissue at speeds of up to 5000 ft/min. Because of the light weight of this paper, it is necessary to make many modifications in the equipment shown in Fig. 7.17.

Many new methods of forming a sheet of paper are being developed. The most common concept is to introduce the pulp slurry between two converging wires and remove the water from both sides. This is termed "twin-wire" forming, and many variations are now in operation. This technique has many advantages and may eventually replace the fourdrinier.

A second method of forming paper uses the cylinder machine. Actually these machines are used primarily to form the multi-ply board used in packaging such as cereal boxes. The cylinder wet end consists of one or more (up to eight) cylinder vats, each of which forms a separate wet web of fibers. Each vat contains a woven wire-covered cylinder rotating in the diluted pulp slurry. The liquid head on the outside of the cylinder is greater than that on

Fig. 7.19. Dry end of large fourdrinier paper machine. (*Courtesy St. Regis Paper Co.*)

the inside, resulting in a flow of water through the wire and out of the vat. The pulp mat that is formed on the face of the cylinder is removed by an endless woolen felt that moves in contact with the cylinder by means of a rubber roll riding on top of the felt. With machines having more than one cylinder, the same felt moves from one cylinder to the next, and the wet sheets from each cylinder are laminated to each other on the bottom side of the felt. Thus, very heavy papers or boards can be fabricated by multivat machines. Because each vat can be supplied with a different type of pulp, it is possible to make boards with a white surface of high-quality pulp and a center of low-cost pulp from waste newspaper or other cheap grades. Because of hydrodynamic problems, cylinder-machine speeds are limited to about 250 ft/min and widths of about 150 inches. However, because of the heavyweight board produced and the widespread use of cheap waste paper for most of the furnish, the cylinder machine is widely used. New forming units looking like miniature fourdrinier units (Ultra-former, Inverformer) have been developed and are rapidly replacing the old-fashioned cylinder vat because their speed is not so limited.

From the wet end of the machine, the wet sheet is conveyed by woolen felts through a series of roll-type presses for further water removal, increasing the consistency to about 35 percent. The sheet then is threaded through the dryer section, consisting of a long series of steam-heated cast iron cylinders that reduce the moisture content to approximately 5 percent, which is about the equilibrium moisture content for cellulosic fibers at 40 to 50 percent relative humidity. Tissue machines use one large dryer (called a Yankee dryer), ranging from 8 to 18 feet in diameter. Because of the light weight of the tissue paper, it can be dried at high speeds on a single dryer.

After drying, the paper is compacted and smoothed by passing through a calender stack consisting of a vertical row of highly polished cast iron rolls. Then the paper is wound into rolls on the reel, as shown in Fig. 7.19.

Finishing and Converting

The objectives in the final stage of the total papermaking system are to improve the paper surface, to reduce rolls and sheets in size, to modify paper for special properties, such as

coat or emboss, to convert to finished products, such as bags and corrugated boxes, and to package for shipping.

Corrugated boxboard is made by gluing sheets of linerboard to each side of a fluted sheet of the corrugating medium. Papers for packing may be laminated to polyethylene film or aluminum foil, or coated with waxes and hot-melt resins. The printing and bag- and box-making industries depend on the production of many mills that produce the several hundred grades of paper used in the United States, and each user may require special paper characteristics to match its process.

The quality of many papers is improved by a surface treatment. A size press about two-thirds of the way along the dryer section can apply a solution of starch to improve surface bonding. Also, paper generally is coated to improve printing properties. A surface coating of a pigment (usually kaolin or china clay, calcium carbonate, or titanium dioxide) and an adhesive (such as starch, casein, and others) is applied to the partially dried web by brush, blade, spray, or other method, and dried during the papermaking (on-machine) operation or in a separate operation. The paper surface is brought to a high finish by passage through the calenders referred to above or through supercalenders. Supercalenders are stacks of alternate steel and densified fiber rolls that create a rubbing action on the sheet, imparting an extra-high gloss to the sheet surface.

Environmental Protection

The manufacture of pulp and paper is a chemical-process industry and produces air emissions, effluents, and solid and toxic wastes that are potential hazards. The paper industry uses large volumes of water as a fiber carrier and as a solvent. An increasing volume of water is being recycled, but makeup water still is required to cover losses. A bleached-pulp and paper mill may use 100 cubic meters (26,400 gallons) of fresh water per metric ton of product and 50,000 cubic meters (13.2 million gallons) daily for a plant producing 500 metric tons of products. In addition to this aqueous effluent that the mill must clean up, it must contend with polluted air and solid and toxic wastes. The nature of these emanations, their sources, and their treatments are summarized in Table 7.8. In this connection the paper industry generally has been in good compliance with governmental environmental regulations although at considerable non-reimbursed capital expense, which amounts to about 10 percent of the cost of the mill.

Using a revolving cylinder or other equipment (save-alls), in the 1930s mills recovered

TABLE 7.8 Summary of Pollution from Pulp and Paper Mill Operations

Type of Pollutant	*Mill Operation*	*Treatment*
Effluents		
Suspended solids (SS): fiber fragments, inerts, clay	Papermaking	Primary—clarification
Pulping residuals	Pulping	Secondary—biological treatment and clarification
Air emissions		
Total reduced sulfur (TRS)	Kraft liquor recovery	Oxidation, precipitation, scrubbing, incineration
Sulfur dioxide	Sulfite pulping	
Particulates	Steam generation	Precipitation, scrubbing
Solid wastes	Effluent treatment	Landfill, utilization, incineration
Toxic wastes: chlorinated compounds	Bleaching	Lime pretreatment, oxidation, biological

for reuse fibers and clay from the paper machine water (white water) system. In the 1970s, procedures for removal of the fibers and clay from the paper mill effluent were incorporated through settling or clarification or primary effluent treatment. About the same time, secondary effluent treatment (biochemical treatment) of the pulp mill effluent was necessary to remove pulping residuals. The purpose of this treatment is to reduce the biological oxygen demand (BOD) of the effluent, which, if untreated, reduces the oxygen content of the stream to a level incapable of supporting aquatic life. The most common method uses microorganisms that react with the wood sugars and other oxygen-consuming compounds in the spent liquors; this is called the activated-sludge method. The products of primary and secondary treatments are sludges, the handling of which is discussed later under solid wastes.

Two objectionable air emissions have characterized pulp mills for years: the sulfur dioxide of the sulfite pulping mill and the malodorous reduced sulfur compounds (TRS) (mercaptans and hydrogen sulfide) of the kraft mill. Still another less noxious air emission is the particulate matter from steam boilers. Coal-burning boilers also emit sulfur dioxide, as is well known. The treatments for the particulate emissions are shown in Table 7.8.

Solid wastes represent the ultimate in mill residues and include the accumulated refuse of the mill and the sludges from primary and secondary effluent treatment. There is difficulty in removing water from the secondary sludge; the primary and secondary sludges often are mixed to aid in water removal, which is important if the sludge is to be incinerated for disposal. The sludges from pulp and paper mills are handled mostly as landfill, and sometimes, if not toxic, they are spread for agricultural purposes.

Most mill solids are slightly toxic, predominantly from chlorination compounds in the wash waters from bleaching. This toxicity can be reduced with lime pretreatment and biological treatment. Toxicity has been the main concern of governmental regulating bodies in recent years.

BOARD AND STRUCTURAL MATERIALS

Board, sheets, panels, and other structural materials are manufactured from wood fibers and various other vegetable fibers, from wood particles, and from paper. The industries making these products are not generally classified as chemical industries; nevertheless, they are closely related to the chemical industry. Fiberboard manufacture is similar to paper-making; particleboards and paper laminates involve the use of synthetic resins and, therefore, chemical technology.

Fiberboard

Fiberboard is the name for rigid or semirigid sheet materials of widely varying densities and thicknesses manufactured from wood or other vegetable fibers. The board is formed by the felting of the fibers from a water slurry or an air suspension to produce a mat. Bonding agents may be incorporated to increase the strength, and other materials may be added to give special properties, such as resistance to moisture, fire, or decay.

Fiberboards are manufactured primarily for panels, insulation, and cover materials in buildings and other structures where flat sheets of moderate strength and/or insulating capacity are required. They also are used as components in doors, cupboards, cabinets, furniture, and millwork.

Classification of fiberboards is best done on the basis of density, as there is a great deal of overlap when classifying by use only. Table 7.9 shows the density classification of fiberboards as well as some of their major uses. Fiberboard factories equalize the mat of loose fibers between screens and rollers, whereas hardboard is press-bonded between hot plates to make it dense and strong. The dividing line between the two types of panels lies at a density of 0.5 g/cm^3. Both are manufactured for many specific uses, and they vary accordingly.

The production of fiberboards goes back to 1898 when the first plant was built in Great Britain. However, large-scale production, mainly of insulation board, developed in the

TABLE 7.9 Classification and Uses of Fiberboards*

Fiberboards	*Density Classification* g/cm^3	*lb/cu ft*	*Major Uses*
Noncompressed (insulation board) semirigid insulation	0.02–0.15	1.25– 9.5	Heat insulation as blankets and batts; industrial cushioning
Rigid insulation board (includes wallboard and softboard)	0.15–0.40	9.5 –25.0	Heat and sound insulation as sheathing, interior panelling, base for plaster or siding, thick laminated sheets for structural decking, cores for doors and partitions, acoustical ceilings
Compressed Intermediate or medium density fiberboard (includes laminated paperboards and homogeneous boards)	0.40–0.80	25–50	Structural use and heat insulation as sheathing base for plaster and siding, interior panelling, containers, underflooring
Hardboard	0.80–1.20	50–75	Panelling, counter tops, components in doors, cabinets, cupboards, furniture, containers, and millwork, concrete forms, flooring
Densified hardboard (superhardboard)	1.20–1.45	75–90	Electrical instrument panels, templets, jigs, die stock

*From information in Fiberboard and Particleboard," Food and Agriculture Organization of the United Nations, Rome, 1958.

United States between the two world wars. The United States is still the largest-producing country and accounts for about one-third of the world output. There has been a much more rapid increase in the production of compressed fiberboards (hardboards) than noncompressed fiberboards (insulation board) during recent years. In 1950 consumption of hardboard was, by weight, roughly 50 percent of insulation board; by 1976 the percentages were almost exactly reversed. Although the total use of insulation board has increased over the past 25 years by 73 percent, per capita consumption has remained constant or even declined, but for hardboard it has doubled. These figures indicate the faster growth rate of the hardboard industry, brought about because favorable prices have stimulated substitution of hardboard for plywood and lumber in construction and furniture. About 85 percent of all insulation board is used in construction, but because it is a nonstructural material, it is not readily substituted for either plywood or hardboard.[23]

Wood is the principal raw material for the manufacture of fiberboards. The species used are numerous, including both softwoods and hardwoods, and vary from region to region. The wood may come from the harvesting of commercial timber, from species not commonly used for lumber or pulp, or from cull timber, logging and forest management residues, and industrial wood residues. Other fiber raw materials for fiberboard manufacture are bagasse (sugar cane residue after sugar extraction) and waste paper. Only minor amounts of other plant fibers are used.

Wood handling and preparation for fiberboard manufacture is much the same as described for pulp and paper. Wood is debarked and chipped with the same type of equipment. If the chips are to be extracted first for rosin or tannin, then cylinder or drum-type chippers may be used instead of disc chippers.

Fibers for fiberboard are coarser and less refined chemically than those used for paper. Processes are used that bring about fiber separation with a minimum loss of chemical components and a maximum yield. The pulping processes used are generally the following: mechanical, thermomechanical,

semichemical, and explosion methods. The pulping processes are essentially the same as that described for making paper pulp.

A unique explosion process for defibering wood was developed by W. H. Mason. Wood chips, about $\frac{3}{4}$ inch long, prepared in conventional chippers and screened, are subjected to high pressure, in a cylinder, commonly called a gun, about 2 × 6 feet in size, and ejected through a quick-opening valve. The elevated temperature softens the chips, and, upon ejection, they explode into a fluffy mass of fibers and fiber bundles. The process involves thermal plasticization of the lignin, partial hydrolysis, and disintegration by the sudden expansion of the steam within the chip.

About 260 lb of wood chips are fed into the cylinder and steamed to 600 psi for 30 to 60 seconds. Then the pressure is quickly raised to 1000 psi (about 285°C) and held only about 5 seconds before sudden release of the charge into a cyclone. The time of treatment at this high pressure and temperature is critical and depends upon the species of wood, chip size, moisture content, and quality of product desired. The steam is condensed in the cyclone, and the exploded fiber falls into a stock chest where it is mixed with water and pumped through washers, refiners, and screens.

The high temperature to which the chips have been subjected causes appreciable hydrolysis of the hemicelluloses, resulting in a somewhat lower yield of pulp than is obtained by mechanical or thermomechanical pulping. The hydrolysis results in a final board product with an enriched lignin content, about 38 percent compared to about 26 percent in the original softwood. The lignin content of the pulp can be varied by controlling the steaming process.

Board Formation. Pulp prepared by any of the above processes may be used for making insulation board. Mechanical (groundwood) pulp was the first type of pulp used in large-scale production of insulation boards, and is still being used in many plants. Pulps from other sources, such as disc mills, may be admixed with it. Groundwood pulp is not considered satisfactory for hardboard, and most hardboard is made from pulp prepared by the explosion process or by defibering with disc refiners.

Board making is basically similar to papermaking and involves refining, screening, mixing of additives, sheet forming, and drying operations. Pressing also is required for hardboards. The pulp is refined and screened prior to sheet formation.

Sizing agents in amounts up to one percent of the fiber are added to the pulp in mixing chests. Paraffin wax emulsion is commonly used for all types of boards. For insulation boards, rosin, cumarone resin, and asphalt also are used. Often a mixture of rosin and paraffin emulsion is used, with 10 to 25 percent rosin in the mixture. For hardboards, paraffin wax is the most common sizing agent although tall oil derivatives and phenol-formaldehyde resins also are used. The sizing agent is precipitated on the fibers by alum, with careful control of pH; the latter may be between 4.0 and 6.5, according to the conditions.

The strength properties of a fiberboard depend mainly upon the felting characteristics of the individual fibers and to a lesser degree upon their interfiber bonding. The felting or forming process usually is done from a water suspension of the fiber at a consistency of around one percent. This is the wet-felting process. An air-felting or dry-forming process is used in a few U.S. plants for hardboard manufacture.

The wet-felting process generally is carried out in a manner similar to papermaking, that is, in a continuous operation on a fourdrinier machine or on a cylinder machine. The machines move more slowly than in the case of papermaking (5–45 ft/min on the fourdrinier), and a coarser mesh wire is employed. In the cylinder-machine method, a single large vacuum cylinder, 8 to 14 feet in diameter, or two cylinders counter-rotating and forming a two-ply sheet are most commonly used. Further water removal is effected by section boxes and press rolls. The wet sheet is cut to

length and conveyed on rollers from the press sections to a tunnel-type dryer. Then the dried sheets are cut into desired sizes.

A third type of wet-felting is a discontinuous method, known as the deckle-box method. The deckle box consists of a bottomless frame that can be raised or lowered onto a wire screen. A measured quantity of stock sufficient to form one sheet is pumped into the deckle box, and vacuum is applied to the lower side of the screen. After most of the water has drained off, pressure is applied from the top to express more water and compact the sheet, reducing its thickness. The deckleframe then is raised and the sheet conveyed to the dryers.

For the recent air-felting process, defibering usually is done in disc mills with control of the moisture content to give the minimum amount possible, consistent with good defibering conditions. Additional moisture may be removed by preheating the air that conveys the fiber from the refiners to the cyclone. The fiber may be further dried in a tunnel or other type of dryer. Fines are removed by either air classification or screens after the dryer. Wax for sizing is introduced either with the chips or added as a spray before or after passing the disc mills (about 2.5% of the weight of the fiber). Sometimes 0.5 to 5 percent of phenolic resin is added, depending upon the quality of board desired. The fiber-blend is fed to a moving screen by a metering unit through a combined air and mechanical action. The fibers felt as they fall on the screen, and the fiber mat thus formed is precompressed between belts and/or rollers. If the board is to be wet-pressed, water is added by spraying.

After the felting or sheet-forming operation, the subsequent operations differ for insulation board and hardboard (see Fig. 7.20). For insulation board, the sheets are dried without further compression, whereas for hardboard, the sheets are either pressed and dried simultaneously (wet-pressing) or are first dried and then pressed (dry-pressing). Air-felted sheets are pressed directly after forming.

Drying and Pressing. The wet-felted sheets for insulation board or for dry-pressed hardboard, containing 50 to 80 percent water, may be dried by any of three methods: (1) tunnel kilns using racks or carts to support the sheets, (2) steam-platen dryers, and (3) continuous roller dryers of single or multideck arrangement. Most widely used is the continuous roller-type multideck dryer, which has an average length of 150 to 300 feet but may be more than 600 feet long. An average dryer will have eight decks and be 12 feet wide.

The pressing conditions greatly affect the board properties. The conditions of time, temperature, pressure, and moisture content will depend upon the fiber in the board and the product desired. In wet pressing, a typical cycle has three phases and lasts from 6 to 15 minutes. The first is a short high-pressure stage (up to 710 psi) to remove most of the free water and bring the board to the desired

Fig. 7.20. Schematic outline for insulation board and hardboard manufacture. (*From "Fiberboard and Particleboard," Food and Agricultural Organization of the United Nations, Rome, 1958.*)

thickness; the second step serves to remove water vapor and requires most of the time; the third stage is a final short period at high pressure to effect a final "cure" or bonding by plastic flow of the lignin. To secure this fiber-to-fiber bond, a temperature of 185°C must be attained, and temperatures up to 210°C may be used to increase the production rate. In the dry-pressing process the cycles are shorter (1.5–3.5 min), and the temperatures and pressures usually are higher.

Conditioning. After the hardboard has been hot-pressed or has been heat-treated or oil-tempered, the moisture content is well below what will be attained at equilibrium with the atmosphere in normal use. Such very dry boards will change dimensions upon picking up moisture and may warp. It is important, therefore, to humidify the boards under controlled conditions before packaging. The desired equilibrium moisture content (EMC) reached will vary from 5 to 12 percent, depending upon the nature of the board and the general humidity conditions in the region of use.

Most humidifying is done in the chambers or tunnels, kept at 80 to 85 percent R.H. and 38 to 50°C. A lesser-used system of conditioning utilizes water spraying and dipping followed by standing to allow uniform absorption.

Special Treatments. Hardboards often are given a special treatment to improve their strength and resistance to moisture. The heat treatment method has come into wide use; the boards, which are kept apart to permit hot air circulation, are heated in chambers by either batch, continuous, or progressive systems. Typical heating conditions are 5 hours at 155 to 160°C. The heating increases the strength (except impact), sometimes as much as 25 percent, and the water resistance is improved. This operation may replace sizing wholly or in part. Some exothermic reaction takes place in the board, and the heat developed must be removed by the hot circulating air to prevent burning. Probably some chemical condensations occur in the wood fiber, producing an internal resin system, and there is the possibility of some crosslinking of large molecules.

Some hardboard is oil-tempered. A drying oil, such as linseed, tung, perilla, soya, or tall oil, or an alkyd resin, is impregnated into the board, by passing the hot-pressed board through a hot oil bath. About 4 to 8 percent of the oil is absorbed. The board then is heated in a kiln with circulating air at 160 to 170°C for 6 to 9 hours. This treatment hardens the oil as well as bringing about chemical reactions in the fiber, and results in greater strength and moisture resistance.

Various additives may be incorporated into insulation boards and hardboards, or they may be surface-treated to bring about resistance to decay, insects, and fire. Pentachlorophenol and copper pentachlorophenate are commonly used for preservative treatments. The sodium salt of pentachlorophenol is added before sizing and is precipitated onto the fiber along with the size by the alum. Special fire-retarding-paint coatings sometimes are used to give resistance to the spread of flame.

Particleboards

Although wood particleboard was patented in the United States as early as 1905, commercial development of the product usually is credited to either Germany or Czechoslovakia in 1941. Production was restricted by a shortage of gluing resins during most of World War II, but in 1945 the first U.S. plant began operation. Particleboard did not really begin to gain market acceptance until the 1960s. From 1964 to 1979 consumption increased at an average annual rate of 9.5 percent. In terms of total consumption of wood-based panels, however, particleboard held only about 15 percent of the market in 1984.[23]

Particleboards are composed of discrete particles of wood bonded together by a synthetic resin adhesive, most commonly urea-formaldehyde or phenol-formaldehyde. The material is consolidated and the resin cured under heat and pressure. The strength

of the product depends mainly upon the adhesive and not upon fiber felting as in the case of fiberboards, although the size and shape of the particles influence strength properties. They may be fine slivers, coarse slivers, planar shavings, shreds, or flakes, and they are divided into two main groups: (1) hammer-mill-produced particles (slivers and splinters from solid wood residues, feather-like wisps to block-shaped pieces from planar shavings) and (2) cutter-type particles, sometimes termed "engineered" particles (flakes and shreds). The various steps in particleboard manufacture are illustrated in Fig. 7.21.

Hammer-milled particles usually vary appreciably in size. Dry raw material produces greater amounts of fines than green wood. Cutting machines (either cylinder-type or rotating-disc-type) give more uniform particles, with the length dimension in the direction of the grain of the wood. The thickness, size, and shape of particles influence the strength of the board. Boards made from sawdust have the lowest strength properties, hammer-milled particles give boards of intermediate strength, and solid wood, cut to flakes, gives boards of highest strength. The latter often are referred to as waferboards.

In another relatively new type of particleboard, the manufacturers align long strands of wood at the surface for increased strength in the direction of panel length. This oriented-strand board (OSB) and waferboard replaced a great deal of plywood sheathing in the early 1980s.

Particleboards may be made in a wide range of densities. Low-density or insulating types are a comparatively recent development in Central Europe, whereas the high-density-hardboard types are a U.S. development. Most particleboard production is in the middle-density range.

Particleboards most commonly are used as core stock for veneer in furniture and in doors, as interior panels for walls and ceilings, as subflooring, as sheathing and siding, and as components in interior millwork. The dense types are used in the same way as fiberboard hardboard, described above. Both dense particleboards and hardboards, after receiving a surface coating, may be printed with decorative designs.

Particleboard production has increased rapidly, both in the United States and worldwide, in recent years. Until recently the most significant limitation on the market for particleboard was the availability of inexpensive plywood. Now plywood is no longer inexpensive relative to particleboard, and the cost factor would seem to favor continued growth of the particleboard industry. However, the advent of waferboard and oriented-strand board has slowed this growth. These new entrants into the structural-panel market were first manufactured on a large scale in Canada, and much of the Canadian output was initially exported to the United States. Within the last decade, however, U.S. production capacity has increased from 9.8 million to 241 million square meters, and there is no indication that this rapid rate of growth will slow in the near future. It appears that waferboard and OSB and several new variants will become the major growth markets among

Fig. 7.21. Schematic outline for particleboard manufacture. (*From "Fiberboard and Particleboard," Food and Agricultural Organization of the United Nations, Rome, 1958.*)

structural panels, and probably will compete most directly with the more expensive softwood plywood.

Paper-Base Laminates

Paper-base laminates are panels or other laminated assemblies composed of many plies of resin-treated paper molded together under high temperatures and pressure to produce rigid structures that no longer have the characteristics of paper. These products are widely used in the electrical and machine industries for insulators, gears, pulleys, and a multitude of machine parts. They possess high impact strength and toughness, good electrical insulation, and high dimensional stability, and are not subject to corrosion; they also have a dampening effect on sound, eliminating rattling and drumming in steel cars and machinery. Furthermore, they easily can be manipulated into complex shapes and can be drilled, turned, and sawed. With these properties, these products are of great industrial value. They also are used in making trays, light flooring panels, table and counter tops, and many other products employing panels.

Paper-base laminates fall into four major classes: (1) mechanical or structural, (2) electrical, (3) punching, and (4) decorative. Phenolic resins are especially suitable and mostly are used where mechanical strength and resistance to heat, water, and electricity are required. For punching-grade laminates, phenolic resins are specially modified with plasticizers or drying oils to yield laminates having good plasticity and elasticity. For electrical grades the phenolics generally are catalyzed with ammonia, amines, or less conductive catalysts in place of the stronger alkaline catalysts otherwise used.

For decorative laminates, urea and melamine resins are the principal resins used. The melamines are used where translucent, light-colored products with good heat and water resistance are required. Polyesters and some melamines are used for low-pressure laminating, enabling the continuous production of counter and table tops by passing the assembly through a set of rolls and then through a heating chamber.

There are two broad general methods for introducing resins into papers for making paper-base laminates: (1) the beater addition process and (2) impregnation of the paper sheet. In the beater addition process the resin is added to the pulp in the beater and then precipitated on the fiber by alum or an acid. The resin adheres to the pulp fibers, which then are formed into a sheet. This type of paper–plastic combination commonly is called resin-filled paper. These resin-filled papers may be made into flat or shaped preforms. The former are flat sheets; the latter are vacuum-felted to a shape closely conforming to that of the final molded product. Shaped preforms are used for deep forms requiring high strength contours. Little flow is required on molding; hence the paper sheet does not break, and good strength properties are obtained. The use of stretchable cross-creped paper for postformable laminates is a new development.

The more common method of application of the resin is by impregnation of the wet or dry paper sheet (web) with a solution of resin. Such papers are called resin-impregnated papers. Water-soluble and alcohol-soluble types of phenolic resins both are used for paper-base laminates. The former tend to give more brittle but more dimensionally stable products than the latter. Phenolic resins are applied mostly by impregnation of the wet or dry paper sheet, producing resin-impregnated papers, although some progress has been made in the slush stock addition procedures to produce phenolic resin-filled papers.

In the wet-web process for impregnating paper for laminates, the wet sheet of paper on the paper machine is carried through a resin bath while the sheet is supported by a wire. The sheet contains up to 65 percent water, and the amount of resin taken up depends upon the moisture content of the paper, solids content of the resin solution, viscosity of the resin solution, temperature of the bath, machine speed, and pressure of the squeeze-rolls. Only water-soluble or water-dispersible resins can be used in the wet-web process.

Generally, however, phenolic impregnations of paper for laminates and other purposes demanding a high resin content are done with dry paper on off-machine equipment. The dry paper sheet is passed continuously through a resin solution, generally an alcoholic solution, and then moves under and between two metering rolls, after which the paper is dried to remove solvent and to complete the condensation of the thermosetting phenolic resin.

The addition of phenolic resins in the form of emulsions to the slush stock by the beater-addition process also has received considerable study. Papers containing 45 to 55 percent resin, and even as high as 65 percent, have been prepared. The resin is precipitated by adding alum to the beater. Papers prepared in this way are highly plastic and are suitable for low-pressure molding. The phenolic resins can also be used in combination with elastomers, such as GR-S types, neoprene, hycar, and vinyl polymers, to produce laminates of high impact strength and greater elongation under tension.

Lignin also is commercially used as a resin for paper laminates. Its low cost and availability as a by-product in pulp manufacture make it attractive. One of the most successful products is made with lignin recovered from the spent liquor of the soda pulping process. The lignin may be added to the pulp suspension in solution and precipitated directly among the fibers, or added in a preprecipitated form.

Properties of paper-base laminates depend upon both the resin and the paper. In general, the final product has the characteristics of the resin used, provided that over 30 percent is resin. The paper acts as a structural reinforcer, greatly enhancing toughness and tensile, flexural, and impact strength.

Electrical properties of the laminate depend upon both the paper and the resin although the amount of resin absorbed is the most important factor affecting electrical insulating properties. Papers of low power factor, good dielectric strength, and high dielectric constant are necessary for electrical grades of laminates. Other desired paper properties for good paper-base laminates are uniformity, cleanliness (freedom from slime, dirt, and fiber bundles), low finish, neutral pH, freedom from chemicals (bleach residues, etc.), low and uniform moisture content (under 4%), and uniform absorbency. High absorbency is desired when high electrical resistance and minimum water absorption are required, as greater amounts of resin are taken up by high-absorbency papers than by those with a low absorbency. For high-impact-strength punching grades of laminates, a low-absorbency paper is necessary to reduce the amount of resin absorbed.

Polymer-Modified Papers

In addition to paper-base laminates, polymers are combined in many ways with paper to develop new properties or to modify or enhance certain properties. Paper commonly is combined with polymers, such as synthetic resins, elastomers, and plastics in general, to produce products that may be classified as follows: (1) wet-strength paper, (2) special-purpose papers containing large amounts of resins or elastomers (chemical-resistant papers, sandpaper backings, gaskets, imitation leather, shoe parts, wood overlay and honeycomb core papers, air and oil filters, battery separators, etc.), (3) plastic-coated papers, and (4) paper-base laminates. The use of plastics in combination with paper has grown enormously during the past three decades. Large quantities of urea and melamine resins as well as polyethylene are used for coating and heating paper and paperboard.

Many other polymers are used in treating and coating papers, including polyesters, styrene, polyvinyl chloride, their copolymers, and others. Polymer-modified papers thus are very important in the present economy.

Overlaid lumber is a composite of lumber and phenolic resin-treated kraft paper. Similar paper overlays are applied to veneer, plywood, fiberboard, and particleboard. Paper honeycomb cores also are made from phenolic resin-treated kraft paper, which is formed into a honeycomb of different geometrical designs, such as figure eight or hexagonal, in special

machines. Simpler types are made from resin-impregnated corrugated sheets, which can be assembled in several ways. This material permits construction of sandwich panels of light weight and high strength. Kraft papers for wood overlays or for honeycomb cores are treated with phenolic resin in the same ways as described above for paper-base laminates.

Paper overlays have three basic uses: masking, decoration, and structures. Masking overlays are used to cover minor defects and provide a more uniform paintable surface. Such overlays contain 20 to 25 percent phenolic resin based on the weight of the paper. For structural purposes, one or more sheets of paper may be used in the laminate. High-density and medium-density types are produced for plywood. The high-density type contains not less than 40 percent of a thermosetting resin, phenolic, or melamine; it has a hard smooth surface not requiring further finishing and may be used for exterior service. The curing of the resin is completed at the same time that the paper is bonded to the wood material in a hot press or, in the case of plywood, at the same time that the veneer is assembled into plywood. The overlay sheet swells and shrinks much less than wood and thus exerts a resistance to the dimensional changes of the wood and may reduce lateral swelling by as much as 40 percent. Overlay papers also upgrade the appearance of low-grade lumber, increase strength properties, improve finishability, and increase resistance to weathering.

Decorative surfaces are obtained by applying a top sheet of white paper on which there is a printed design. This is covered with a clear coating of melamine or vinyl resin. Or, a thin transparent (when cured) paper impregnated with melamine resin may be applied to a decorative veneer, providing a permanent protective finish.

For honeycomb cores, either water- or alcohol-soluble phenolic resins may be used. Many types of facings may be glued to the honeycomb cores: veneers, plywood, hardboards, asbestos board, aluminum, stainless steel, and paper-plastic laminates. Thin sheet material may be used because of the almost continuous support of the core.

The honeycomb sandwich possesses great strength in relation to its weight. It may carry loads as much as 25 tons/ft^2. Its strength and weight vary with the weight of the paper, quantity of resin impregnated, and honeycomb design.

Honeycomb sandwich construction has many uses, such as in airplanes, cargo containers, truck and trailer bodies, railway passenger cars, cabins, barns, airplane hangars, house floors, walls, roofs, doors, and a variety of other products. Besides combining strength with lightness, honeycomb sandwich material has high rigidity, good insulation properties, resistance to fungi and pests, and durability to temperature extremes.

Modified Wood and Wood Composites

Wood has a number of disadvantageous properties as described earlier in this chapter. It is biodegradable, flammable, dimensionally unstable to water, and degradable by ultraviolet light, acids, and bases. There has been a considerable amount of research expended over the years to correct these defective properties for both solid wood and reconstituted wood products.

A wood product of improved dimensional stability can be produced through heat treatment of wood at temperatures varying from 150°C to 320°C. For good temperature control and to avoid strength loss the heat treatment preferably is performed under the surface of molten metal or a fused salt for exclusion of oxygen. The modified wood product is referred to as staybwood.

Treatment of wood with polyethylene glycol (PEG) is the most commonly used method to impart dimensional stability to wood. PEG in a water solution can be used to permeate the lumens and penetrate the cell wall. A 30 to 50 percent solution of PEG in water generally is used on green wood; no pressure is required. Treatment times are adjusted so that a uniform uptake of 25 to 30 percent of PEG is achieved, based on the dry weight of the wood; temperatures range from

21°C to 60°C. The wood may require weeks of soaking, depending on species, density, and thickness. After treatment, the wood is dried in a well-ventilated space, preferably heated.

The PEG is not "fixed" or cured in the wood, and it leaches out if the wood contacts water. For this reason, PEG-treated wood usually is finished with a penetrating oil or polyurethane varnish. PEG is used primarily to treat cross sections of trees for plaques, clock faces, and tabletops. It also is used to stabilize bowls and other turnings, carvings, and rifle stocks, and to treat archaeological water-logged wood.

Resin impregnation can be used to add bulking agents to the cell wall. These resins have a low degree of polymerization, or the monomers are polymerized after impregnation. The best fiber-penetrating thermosetting resins are highly water-soluble, phenol-formaldehyde, resin-forming systems with low initial molecular weights. Green or dry veneers usually are treated by soaking or pressure treatment with a water-soluble resin in a 30 percent aqueous solution. Following impregnation, the wood is dried slowly and then heated at about 150°C for 20 minutes to set the resin. Laminates of treated wood are constructed by gluing the individual sheets together. This product, called Impreg, contains 25 to 35 percent resin and has a density about 20 percent higher than the original wood and an anti-shrink efficiently (ASE) of about 65 percent. As with PEG, the water-soluble resin penetrates the cell wall and keeps the wood in a swollen state. Unlike PEG, the phenol-formaldehyde resin is cured by heating to form a high-molecular-weight, water-insoluble resin in the cell wall. Thus, Impreg can be used in contact with water. Its compressive strength is higher than that of untreated wood, but it has much less toughness. It shows considerable resistance to decay and termite and marine-borer attack. Treating with phenolic resins imparts high acid resistance, greatly increases the temperature to which wood can be heated without appreciable change in properties, and increases electrical resistance. The largest industrial application of Impreg is in die molds for automobile body parts and other uses. Despite its high cost, the dimensional stability and plasticity of Impreg make it commercially viable.

Compreg is a highly dense product obtained by applying pressure to dry, resin-treated veneers during heat cure. It offers most of the advantages of Impreg. Its mechanical properties are appreciably better than those of the original uncompressed wood because of a two- to three-fold increase in density ($1350\,kg/m^3$). The strength properties of Compreg are increased in proportion to the degree of compression. Only its toughness is lower than that of untreated wood, although greater than that of Impreg. Because of the plasticizing action of the resin-forming materials, treated wood can be compressed under considerably lower pressures than dry, untreated wood. For example, when subjected to a pressure of only 1.7 MPa (244 psi) at 149°C, treated spruce, cottonwood, and aspen wood, dried to a moisture content of about 6 percent but not cured, are compressed to about half the original thickness and a specific gravity of about 1.0.

In a 24-hour water-soaking test, Compreg has an ASE value of 95 percent. The rate of water pickup is so slow that complete swelling equilibrium of a 13-mm specimen is not reached after one year at room temperature. The product is brown and can be buffed to a high polish. Compreg is produced commercially in small quantities and is used for knife handles, gears, certain musical instruments, and decorative articles.

Composites of wood with vinyl polymers have been developed. Because the vinyl polymers are clear, colorless, thermoplastic materials, they do not significantly discolor the wood; thus its natural beauty is retained, whereas the phenolic resins darken the wood. However, the hygroscopic characteristics of the wood substance are unaffected because little, if any, resin penetrates the cell walls; the ASE is only about 10 to 15 percent. The high resin content (70–100%) greatly reduces the normally high void volume of wood, however. Thus, the elimination of this important pathway for vapor diffusion slows down the response of the wood substance to changes in

relative humidity, and its moisture resistance is greatly improved compared to the original wood. The hardness of the wood-plastic composite also is greatly improved.

A variety of vinyl monomers, such as methyl methacrylate and styrene, may be used. Complete filling of the cell lumens and other voids (the "full-cell process") is easily accomplished by first subjecting the wood to a partial vacuum (about 0.3 in. of Hg) and then covering it with the monomer and soaking it for 2 to 6 hours, depending upon the species of wood and its dimensions. Some penetration of the monomer into the cell walls also may be obtained by using a diffusion process, such as a solvent-exchange method.

Polymerization of the vinyl monomer in the wood may be done with either radiation or free radical catalysts. The polymerization of the vinyl monomers in both processes depends upon the same mechanism, that is, initiation by free radicals. In the radiation process, the gamma rays passing through the monomer and the woody tissue create a large number of excited and ionized molecules, many of which break into fragments, namely, organic free radials ($R^{\cdot}$). These act as the initiator for the polymerization of an unsaturated monomer. Alternatively, the free radicals may be formed by thermal decomposition of compounds involving a weak bond. Commercially, the catalyst 2,2′-azobisisobutyronitrile now is most widely used, as it forms free radicals at a lower temperature than benzoyl peroxide.[24]

If the end use of the wood–polymer composite requires an abrasive (sanding) or cutting process that brings about high temperatures, the thermoplastic polymer will melt, causing machining difficulties. To prevent such melting, a crosslinking substance such as diethylene glycol dimethacrylate is added to the monomer before impregnation into the wood (about 5% of the volume of the monomer). Wood–plastic materials are used in parquet flooring, certain sporting equipment, musical instruments, and decorative articles.

Wood also can be modified by chemical reaction directly with the hydroxyl groups present in the polymer constituents in the cell wall of the wood. This treatment confers bulk to the cell wall with the help of a permanently bonded chemical. Many reactive chemicals have been used to modify wood chemically. For best results, the chemicals should be capable of reacting with the wood hydroxyl groups under neutral or mildly alkaline conditions at temperatures below 120°C. The chemical system should be simple and must be capable of swelling the wood structure to facilitate penetration. The complete molecule must react quickly with wood components, yielding stable chemical bonds; the treated wood still must possess the desirable properties of untreated wood. Anhydrides, epoxides, and isocyanates best conform to these conditions, conferring ASE values of 60 to 75 percent at chemical weight gains of 20 to 30 percent. Reactions of the chemicals with wood, as shown below, are fast and complete, and stable chemical bonds are formed. The systems swell the wood and penetrate well.[25]

Epoxides:

Isocyanates:

Anhydrides:

In the epoxide reaction, a new hydroxyl group originating from the epoxide is formed; and from this new hydroxyl, a polymer can begin to form. Given the ionic nature of the

reaction and the availability of alkoxyl ions in the wood components, the chain length probably is short because of chain transfer.

The optimum reaction conditions for reaction with epoxides are a 95/5 (v/v) mix of epoxide with triethylamine at 120°C and 150 psi nitrogen. Depending on the isocyanate used, reaction conditions of 120°C and 150 psi with no catalyst and 5 percent triethylamine or 5 to 35 percent dimethylformamide are optimum.

Of several anhydrides studied, acetic anhydride reacted the most readily. Reactions were carried out by refluxing the wood in a xylene/acetic anhydride solution or with acetic anhydride vapors alone at 120°C. With this system, for each mole of acetate bonded onto the wood a mole of acetic acid is generated as a by-product. Although this by-product generation is a disadvantage of the process, the chemical system does penetrate and react quickly with wood, without a catalyst. It is not so sensitive to moisture as are the epoxide and isocyanate systems.

Recently a very facile dip acetylation procedure has been developed for wood chips and fibers. The wood is dried at 105°C and then dipped in acetic anhydride for one minute and drained. Then the sample is placed in a preheated (120°C) reactor for different lengths of time, depending on the degree of substitution of acetyl groups desired in the sample.[26] The swelling of aspen flakeboard made from such chemically modified flakes is dramatically reduced. For example, while unmodified flakeboard swelled by over 60 percent after immersion in water for 5 days, the modified material swelled by only about 8 to 22 percent, depending on the extent of chemical modification.

Molded and Formed Wood. Historically wood has been bent and shaped first by steaming, which plasticizes the wood polymers. The new shape then is set in place by cooling and drying in the desired conformation. However, the curvature possible by this method is limited, and the process is time-consuming. Some chair backs and solid hickory and ash skis still are steam-bent.

Anhydrous ammonia also is known to cause temporary platicization of wood. The ammonia swells and plasicizes both the lignin and the cellulose, and the crystalline structure of the cellulose is converted to a different form in the process. To shape the wood, it is immersed in liquid ammonia or treated with gaseous ammonia under pressure until the cell walls have been penetrated and the wood becomes pliable and flexible. In this condition it is easily shaped and formed by hand or mechanically. The ammonia readily vaporizes and evaporates from the wood, so that the wood regains its normal stiffness but retains the new form into which it has been shaped. With this process the wood can be distorted into quite complex shapes without springing back to its original form. Treating plants have been developed on a pilot-plant scale, but the process has not been widely adopted.

There has been a concerted research effort in recent years to develop molded wood-based material. The Japanese have emphasized chemical modification approaches and have made considerable progress toward the development of a thermoplastic wood-based material.[27] Their most promising results are based on esterification of wood meals with subsequent thermomolding. As the size of the aliphatic group was increased, the melting temperature of the modified wood under pressure was decreased. It also was found that a very small degree of additional grafting of polystyrene to the esterified wood resulted in a dramatic improvement in the thermoplasticity of the product. Matsuda[28] developed a very facile method for production of moldable wood by simply heating the wood meal with an acid anhydride for 3 hours to temperatures greater than 60°C. The wood meal was readily moldable at high temperatures (180°C) under pressure (570 kg/cm^2) for 10 minutes. The moldability of the esterified wood decreased in the following order: succinic anhydride > maleic anhydride > phthalic anhydride.

Although it is possible to make wood thermoplastic through chemical modification, the handling of the finely divided wood (wood meal) becomes a problem when more complex

shapes or deep draws are desired. Therefore, a totally new approach for the creation of molded products based on wood was developed by Hunter Brooks, formerly of the Triad Corp., Detroit, Michigan. The raw material is a dry attrition-milled wood fiber (80%), a polyester staple fiber 1 to 2 inches long (10%), and a phenolic resin (10%). The materials are formed into a nonwoven web on a Rando-webber machine. The longer polyester fibers are incorporated so that the web can be needle-punched to form a nonwoven fabric with good mechanical integrity. The phenolic resin is applied to the web either as a dry powder or as an aqueous solution. This fiber composite then is easily compression-molded into any of a variety of shapes. For deep-draw compression molding, patterns may be cut into the nonwoven web to allow the fiber mat to drop deep into the mold. Any type of fibrous material, synthetic or natural lignocellulosic, can be utilized in the web; so the process offers considerable flexibility for production of a wide variety of products engineered to specific applications. The processing equipment developed by Hunter Brooks has been donated to the USDA Forest Products Laboratory, where an active research program in cooperation with the University of Wisconsin is under way to perfect the process for industrial applications.[29]

A major application of these types of molded products would be for interior uses in automobiles, such as head liners, door panels, and dashboards. Although this is a low-cost, low-performance application, it represents a very large-volume market. Indeed, wood is already utilized in applications of this type, but as a finely ground flour that serves as a filler (up to 40%) in extrusion-molded polyolefin products. The use of recycled fiber in this process and the one described above offers the potential of even greater cost reductions, combined with alleviation of solid waste disposable problems.

PRESERVATIVE TREATMENT OF WOOD

Wood, as a natural plant tissue, is subject to attack by fungi, insects, and marine borers. Some species of wood are more resistant to decay than others (e.g., the heartwood of cedars, cypress, and redwood) because of the presence of natural toxic substances among the extractable components. Most woods, however, are rapidly attacked when used in contact with soil or water, or when exposed to high relative humidities without adequate air circulation. Wood for such service conditions requires chemical treatment with toxic chemicals, collectively termed wood preservatives. The service life of wood may be increased 5- to 15-fold, depending upon the conditions of preservative treatment and the nature of the service.

The preservative treatment of wood is the second largest chemical wood-processing industry; pulp and paper manufacture is the most important. Since 1950, the average annual volume of wood treated has been 250 to 300 million cubic feet. The more important types of wood products treated are shown in Fig. 7.22.

Preservative Chemicals

Owing to their toxic nature, all of the commercial wood preservatives presently used in the United States are effective in preventing attack by microorganisms. However, because of concern that these chemicals will have a deleterious effect on the environment, alternative methods based on nontoxic procedures also are being investigated. These nonconventional approaches are based on chemical modification, and are described briefly at the end of this section.

Toxic chemicals used for the preservation of wood may be classified as follows:

- *Organic liquids of low volatility and limited water solubility*:
  Coal-tar creosote
  Creosote–coal tar solutions
  Creosote–petroleum solutions
  Other creosotes
- *Chemicals dissolved in organic solvents, usually hydrocarbons*:
  Chlorinated phenols (principally pentachlorophenol)
  Copper naphthenate
  Solubilized copper 8-quinolinolate

Fig. 7.22. Relative distribution of wood products treated with preservatives.

- *Water-soluble inorganic salts*:
  Acid copper chromate
  Ammoniacal copper arsenite
  Chromated copper arsenate
  Chromated zinc chloride
  Fluor chrome arsenate phenol

Creosote from coal tar is the most widely used wood preservative for several reasons: (1) it is highly toxic to wood-destroying organisms; (2) it has a high degree of permanence due to its relative insolubility in water and its low volatility; (3) it is easily applied, with deep penetration easily obtained; and (4) it is relatively cheap and widely available. For general outdoor service in structural timbers, poles, posts, piling, and mine props, and for marine uses, coal-tar creosote is the best and most important preservative. Because of its odor, dark color, and the fact that creosote-treated wood usually cannot be painted, creosote is unsuitable for finished lumber and for interior use.

Coal-tar creosote is a mixture of aromatic hydrocarbons containing appreciable amounts of tar acids and bases (up to about 5% of each), and has a boiling range between 200 and 355°C. The important hydrocarbons present include fluorene, anthracene, phenanthrene, and some naphthalene. The tar acids are mainly phenols, creosols, xylenols, and naphthols; the tar bases consist of pyridines, quinolines, and acridines.

Often coal tar or petroleum oil is mixed with coal-tar creosote, in amounts up to 50 percent, as a means of lowering preservative costs. Because coal tar and petroleum have a low toxicity, their mixtures with creosote are less toxic than is creosote alone.

A number of phenols, especially chlorinated phenols and certain metal-organic compounds, such as copper naphthenate and phenyl mercury oleate, are effective preservatives. Pentachlorophenol and copper naphthenate are most commonly used, and are carried into the wood in 1 to 5 percent solutions in petroleum oil. Pentachlorophenol is colorless, and can be applied in clear volatile mineral oils to millwork and window sash requiring a clean, nonswelling, and paintable treatment.

Inorganic salts are employed in preservative treatment where the wood will not be in contact with the ground or water, such as for indoor use or where the treated wood requires painting. They are also satisfactory for outdoor use in relatively dry regions.

Preservation Process

The methods for applying preservatives to wood are classified as follows:

- *Nonpressure processes*:
  Surface (superficial) applications by brushing, spraying, or dipping
  Soaking, steeping, and diffusion processes
  Thermal process
  Vacuum processes
  Miscellaneous processes

- *Pressure processes*:
  Full-cell process (Bethell)
  Empty-cell processes (Rueping and Lowry)

Brush and spray treatments usually give only limited protection because the penetration or depth of capillary absorption is slight. Dip treatments give slightly better protection. Organic chemicals dissolved in clear petroleum solvents often are applied to window sash and similar products by a dip treatment of 1 to 3 minutes.

Cold soaking of seasoned wood in low-viscosity preservative oil for several hours or days and the steeping of green or seasoned wood in waterborne preservatives for several days are methods sometimes employed for posts, lumber, and timbers on a limited basis. The diffusion process employs water-borne preservatives that will diffuse out of the treating solution into the water in green or wet wood.

The most effective of the nonpressure processes is the thermal method of applying coal-tar creosote or other oil-soluble preservatives, such as pentachlorophenol solution. The wood is heated in the preservative liquid in an open tank for several hours, after which it is quickly submerged in cold preservative in which it is allowed to remain for several hours. This is accomplished either by transferring the wood at the proper time from the hot tank to the cold tank, or by draining the hot preservative and quickly refilling the tank with cooler preservative. During the hot treatment, the air in the wood expands, and some is expelled. Heating also lowers the viscosity of the preservative so that there is better penetration. When the cooling takes place, the remaining air in the wood contracts, creating a partial vacuum that draws the preservative into the wood. For coal-tar creosote, the hot bath is at 210 to 235°F, and the cold bath at about 100°F. This temperature is required to keep the preservative fluid.

The hot- and cold-bath process is widely used for treating poles and, to a lesser extent, for fence posts, lumber, and timbers. This process gives the most effective results of the common nonpressure processes, which most nearly approach those obtained by the pressure processes.

The vacuum processes involve putting the wood under a vacuum to draw out part of the air. The wood may be subjected to a vacuum alone or to steaming and a vacuum before being submerged in a cold preservative. These methods are used to a limited extent in the treatment of lumber, timber, and millwork.

Commercial treatment of wood is most commonly done by one of the pressure processes, as they give deeper penetrations and more positive results than any of the nonpressure methods. The wood, on steel cars, is run into a long horizontal cylinder, which is closed and filled with preservative. Pressure is applied, forcing the preservative into the wood.

There are two types of pressure treatment, the full-cell and the empty-cell. The full-cell process seeks to fill the cell lumens of the wood with the preservative liquid, giving retention of a maximum quantity of preservative. The empty-cell process seeks deep penetration with a relatively low net retention of preservative by forcing out the bulk liquid in the wood cells, leaving the internal capillary structure coated with preservative.

In the full-cell process, the wood in the cylinder first is subjected to a vacuum of not less than 22 in. Hg for 15 to 60 minutes, to remove as much air as possible from the wood. The cylinder then is filled with hot treating liquid without admitting air. The maximum temperature for creosote and its solutions is 210°F, and for water-borne preservatives it is 120 to 150°F, depending upon the preservative. Then the liquid is placed under a pressure of 125 to 200 psi, and the temperature and pressure are maintained for the desired length of time, usually several hours. After the liquid is drawn from the cylinder, a short vacuum is applied to free the charge of surface-dripping preservative.

In the empty-cell process, the preservative liquid is forced under pressure into the wood, containing either its normal air content (Lowry process) or an excess of air, by first

subjecting the wood to air pressure before applying the preservative under pressure (Rueping process). In the former case, the preservative is put in the cylinder containing the wood at atmospheric pressure, and, in the latter case, under air pressure of 25 to 100 psi. After the wood has been subjected to the hot preservative (about 190–200°F) under pressure (100–200 psi in the Lowry process and 150–200 psi in the Rueping process) and the pressure has been released, the back pressure of the compressed air in the wood forces out the free liquid from the wood. As much as 20 to 60 percent of the injected preservative may be recovered, yet good depth of penetration of the preservative is achieved.

Preservative Retention

Retention of preservative generally is specified in terms of the weight of preservative per cubic foot of wood, based on the total weight of preservative retained and the total volume of wood treated in a charge. Penetration and retention vary widely between different species of wood, as well as with woods of the same species grown in different areas. In most species, heartwood is much more difficult to penetrate than sapwood. Also, within each annual growth ring there is variability in penetration, the latewood generally being more easily treated than the earlywood.

The American Wood-Preservers' Association Standards specify methods of analysis to determine penetration and retention. They also specify minimum retention amounts for different preservatives according to the commodity, the species, the pretreatment of the wood, such as kiln drying, and the end use of the commodity. Heavier retention is required for products in contact with the ground (poles, timbers, etc.) or with marine waters (piles, timbers, etc.). Unprotected wood in contact with the ground is subject to severe attack by fungi and insects, and, in contact with seawater, it is quickly destroyed by marine borers. For wood products to be used in contact with the ground or marine waters, creosote is the major preservative employed because it can be readily impregnated to give high retention and good protection, and it is not leached out by water.

Nonconventional Wood Preservation

Chemical modification as a possible preservative treatment for wood is based on the theory that enzymes (cellulose) must directly contact the substrate (cellulose), and this substrate must have a specific configuration. If the cellulosic substrate is chemically changed, this highly selective reaction cannot take place. Chemical modification also can change the hydrophilic nature of wood. In some cases water, a necessity for decay organisms, is excluded from biological sites. The chemicals used for modification need not be toxic to the organism because their action renders the substrate unrecognizable as a food source to support microbial growth. For wood preservation, this means that it is possible to treat wood in such a manner that attack by wood-destroying fungi will be prevented, and the material will be safe for humans to handle. For applications of wood in which human contact is essential, nontoxic preservatives may well be specified or required in the future. An added benefit to wood of most chemical modification treatments is that the resulting bulking action gives the treated wood very good dimensional stability. The chemical modification of wood involves a chemical reaction between the hydroxyl groups in wood and a simple single chemical reagent, with or without a catalyst, to form a covalent bond between the two, as previously described for dimensional stability treatments.[30] The wood component may be cellulose, hemicellulose, or lignin.

The treated wood must continue to possess the desirable properties of untreated wood: the strength must remain high, there must be little or no color change (unless a color change is desirable), and it must be a good electrical insulator, safe to handle, gluable, paintable, and so on. These chemicals, once reacted, are effective in preventing attack by microorganisms, but they are not toxic to the decay organisms. The important factor in preventing attack is to attain a treatment level that

inhibits the growth of the organisms. A recent review on this subject shows that treatments with acetic anhydride, dimethyl sulfate, acrylonitrile, butylene oxide, phenyl isocyanate, and β-propiolactone all give good rot resistance at a 17 to 25 weight percent gain (WPG).[30] The decay resistance of acetylated wood is proportional to the WPG, and the degree of dimensional stability also is proportional to the WPG; so the exclusion of cell wall or biological water may be a very important factor in the decay resistance mechanism.[31]

In preliminary tests, alkylene oxide–treated southern pine was found to be resistant to termite attack and attack from the marine borers *Teredo* (shipworm) and *Limnoria*.

In conclusion, chemical modification of wood results in a treatment that is nontoxic, effective, and nonleachable. The high chemical treatment level required for effectiveness, however, results in a rather expensive treatment. Dimensional stability also is obtained at these high (17–25 WPG) substitution levels; so, for those products where both not resistance and dimensional stability are important, the present state of the technology is close to a viable industrial process. The use of toxic chemicals that are permanently bound to the wood components by this treatment may be an environmentally acceptable preservation method.

FIRE-RETARDANT TREATMENT OF WOOD

The Romans first treated wood for fire retardancy in the first century A.D. They used solutions of alum and vinegar to protect their boats against fire. In 1820, Gay-Lussac advocated the use of ammonium phosphates and borax for treating cellulosic material. Many of the promising inorganic chemicals used today were identified between 1800 and 1870. Since then, the development of fire retardants for wood has accelerated. Commercially treated wood became available after the U.S. Navy (1895) specified its use in ship construction, and New York City (1899) required its use in buildings over 12 stories tall. Production reached over 65 million board feet in 1943, but by 1964 only 32 million board feet was treated annually.[32]

However, starting in 1979 changes were made in some building codes that allowed structural changes if fire retardant (FR) treated wood was used in place of untreated plywood for roof sheathing. The structural changes resulted in sufficient savings such that FR-treated wood was utilized more extensively. Also the use of FR-treated plywood was mandated at other specific sites such as prisons.

Fire-retardant treatments for wood can be grouped into two general classes: (1) those impregnated into the wood or incorporated into wood composite products and (2) those applied as paint or surface coatings. Chemical impregnation has the greater use, primarily for new materials, whereas coatings have been limited primarily to materials in existing constructions. There are advantages and disadvantages to each class. Coatings are applied easily, and they are economical. Chemical impregnation usually involves full-cell pressure treatment and can be costly. A coating is subject to abrasion or wear that can destroy the effectiveness of the fire retardant. Chemical impregnations deposit the fire retardant within the wood, so that if the surface is abraded, chemicals are still present. On-site application of surface coatings requires strict control of the amount applied to ensure correct loading levels for a particular flame-spread rating. Both coating and impregnation systems are based on the same chemical compounds although the formulations for each vary. Among the most commonly used chemicals for impregnation treatments are diammonium phosphate, ammonium sulfate, borax, boric acid, zinc chloride, and most recently the leach-resistant amino-resin systems. These compounds have different characteristics with respect to fire resistance. Ammonium phosphate, for example, is effective in checking both flaming and glowing; borax is good in checking flaming but is not a satisfactory glow retardant. Boric acid is excellent in stopping glow but not so effective in retarding flaming. Because of these different

characteristics, mixtures of chemicals usually are employed in treating formulations.

Fire-Retardant Formulations

Many chemicals have been evaluated for their effectiveness as fire retardants. Today most fire retardants for wood are based on phosphorus, nitrogen, boron, aluminum trihydrate, and a few other compounds. Phosphorus and nitrogen frequently are used together because they behave synergistically; amino-resins are an example of such a combination.

Most fire-retardant formulations are not resistant to leaching by water. Therefore, there have been increased efforts to develop leach-resistant chemicals that can be impregnated into wood products for use in exterior or high-humidity applications. Some of the proposed leach-resistant systems include chemical combinations that form insoluble complexes, amino-resin systems, and monomers that polymerize in the wood. A common amino-resin system for exterior use is dicyandiamide phosphoric acid formaldehyde. Guanylurea phosphate-boric acid also is a commonly used organic phosphate salt for modern commercial fire-resistant wood.

The American Wood-Preservers' Association Standards specified the four types of fire-retardant formulations given below (A, B, C, D). Many newer formulations have been developed by commercial enterprises and are proprietary.

Type A

Chromated zinc chloride—a mixture of sodium dichromate and zinc chloride having the composition: hexavalent chromium as CrO_3, 20% and zinc as ZnO, 80%.

Type B

Chromated zinc chloride (as above)	80%
Ammonium sulfate	10%
Boric acid	10%

Type C

Diammonium phosphate	10%
Ammonium sulfate	60%
Sodium tetraborate, anhydrous	10%
Boric acid	20%

Type D

Zinc chloride	35%
Ammonium sulfate	35%
Boric acid	25%
Sodium dichromate	5%

Minimum and maximum limits of variation in the percentage of each component in the above formulations are specified in the standards.

The impregnation methods are similar to those employed for the preservative treatment of wood by water-borne salts using pressure processes. The maximum temperature of the solution must not exceed 140°F for formulation Types A, B, and D, and must not be over 160°F for Type C. After treatment, the wood must be dried to remove the water solvent to a moisture content of 19 percent or less. For most uses, the wood is kiln-dried to a moisture content of under 10 percent.

For effective fire protection larger amounts of chemicals must be deposited in the wood than the amounts necessary when water-borne chemicals are used for decay prevention. Whereas for the water-soluble toxic salts retentions from 0.22 to 1.00 lb/ft^3 of wood are specified according to commodity standards in order to give good protection against decay and insects, as much as 5 to 6 lb of some fire retardants may be required for a high degree of effectiveness against fire. Usually, however, smaller amounts will give a good degree of protection. For example, formulation Type B when impregnated in amounts of 1.5 to 3 lb/ft^3 of wood provides combined protection against fire, decay, and insects. Because the amino-resin systems are leach-resistant, high application levels are not necessary in practice.

Fire retardancy of wood involves a complex series of simultaneous chemical reactions, the products of which take part in subsequent reactions. Most fire retardants used for wood increase the dehydration reactions that occur during thermal degradation so that more char and fewer combustible volatiles are produced. The mechanism by which this happens depends on the particular fire retardant and the thermal-physical

environment. The effectiveness of a fire-retardant treatment depends upon the performance rating of the treated material when tested in accordance with ASTM E84 (no greater flame spread than 25).

CONVERSION OF WOOD TO ENERGY, FUELS, AND CHEMICALS

Wood has been a source of energy and chemicals for hundreds of years and continues to be an important raw material for specific chemicals. The use of wood as a primary source of industrial chemicals decreased dramatically in the 1940s when oil became the preferred raw material. The term "silvichemicals" sometimes is used to refer to wood-derived chemicals analogous to petrochemicals.

The use of wood for energy, fuels, and chemicals can be conveniently divided into four major categories: direct combustion, saccharification-fermentation (SF), thermal decomposition, and thermochemical liquefaction. These methods are discussed in more detail in the following sections.[5]

Direct Combustion

The concept of using wood as a source of energy through direct combustion dates back to the very beginning of human existence. As soon as early people learned to use fire, wood became the major source of energy. It is important to note that fuel is still the major use of wood on a worldwide basis, amounting to an estimated $1.2 \times 10^9\,m^3$/year compared to $0.8 \times 10^9\,m^3$/year for lumber and other solid wood products, $0.34 \times 10^9\,m^3$/year for pulp, and $0.2 \times 10^9\,m^3$/year for other uses. Thus, on a global basis, nearly 50 percent of the wood harvest is used directly for fuel. In many nonindustrialized countries, fuel amounts to 80 to 90 percent of total wood consumption, compared with an average of 7 percent for industrialized regions.

The average stored energy content of wood is 8600 Btu/lb. Variations in heat values for wood as a function of species occur within a relatively narrow range of 8000 to 10,000 Btu/lb. Bark values are slightly higher than wood values, about 10,000 Btu/lb.

The greater the oxygen content of a carbon compound or carbonaceous substance is, the lower the heat of combustion. Lignin, with an elemental analysis of $C_{10}H_{11}O_2$ compared to $C_6H_{10}O_5$ for cellulose, has a lower degree of oxidation and a higher heat of combustion than cellulose. The fossil fuels which have very low or no oxygen content have much higher energy contents, for example, 12,000 to 13,500 Btu/lb for coal, 18,000 to 19,900 Btu/lb for oil, and 18,550 Btu/lb for natural gas.

Wood and other biomass materials are hygroscopic and retain substantial amounts of moisture (15–60%). Hence, some of the thermal energy of the wood is lost in vaporizing this water to steam during combustion. In addition, wood has a lower density than coals, so that a greater volume of material must be gathered and transported to provide the same heating value as that of coal. Wood has a negligible sulfur content, less than 0.1 percent compared to much larger amounts in coal (avg. 2%) and petroleum. Sulfur from coal and petroleum creates serious environmental problems by causing acid rain. Also, ash from coal presents a disposal problem and creates unhealthy fly ash, whereas the ash content of wood is less than 0.5 percent. The only environmental hazard produced from burning wood as a fuel is the production of particulates.

Most wood and wood-derived materials (spent pulping liquors) that are used for energy are consumed by the forest products industry itself. The source of this fuel is almost entirely in the wood processing and manufacturing operations, termed "manufacturing residuals" or "mill residues." Forest residues from logging operations and the noncommercial trees in the forest also are utilized for fuels.

Wood wastes supply the forest products industry with roughly 40 to 50 percent of the energy requirements in the form of spent liquors and wood and bark residues. The spent pulping liquor contains 40 to 50 percent of the wood as dissolved lignin and a large

variety of other organic compounds. The spent liquor is used almost totally, after evaporation to about 65 percent solids, to produce steam and to recover the inorganic chemicals used in the pulping operation. The Tomlinson furnace is used almost universally and is the most efficient system for combined energy and chemical recovery yet devised.

The mixed wood and bark residues burned directly are collectively termed "hog fuel." Over the past decade, there has been a trend to channel more and more of the cleaner, drier, and larger-sized mill residues into raw materials for products. The coarse wood residues (slabs, trim, edgings, etc.) are chipped and sent to pulp mills. Also, considerable sawdust now is being pulped instead of being burned for fuel. Some of the residues, including shavings, also are used for particleboard. This means that hog fuel contains the less desirable, dirtier, and wetter forms of waste.

Combustion Technology. The hog fuel boiler represents the conventional technology of using wood for energy to produce process steam. Modern wood-fuel boilers have traveling grates and use mechanical draft fans. Many are of gigantic size, handling as much as 500,000 lb of wood per hour. Spreader stokes distribute the fuel evenly over the large grate areas to ensure efficient combustion. The ash can be removed continuously with traveling grates. Modern instrumentation meters air flow, controls fuel-to-air ratios, and meters combustion efficiency and stack emissions. With these developments, the energy recovery and environmental acceptability of wood-burning systems have been improved in the forest products industry.

The deterioration of hog fuel quality referred to above makes it more difficult to obtain efficient combustion, high reliability, and low stack emissions simultaneously. To overcome this difficulty, the fuel can be processed before combustion to remove dirt and moisture and produce clean-burning, efficient fuel. This can be done by drying, screening, grinding, washing, and pelletizing or densification of the hog fuel in order to remove noncombustible dirt and moisture, reduce the size of large and slow-burning material, and agglomerate small and fine material. Size uniformity is important for ease in transportation, storage, and fuel feeding. Dirt and moisture reduction reduces transportation costs, promotes efficient combustion, and minimizes air emissions. Another approach for using low-quality hog fuel is the development of improved combustors. Examples are: (1) the fluidized-bed combustors, which obtain excellent wood combustion at relatively high efficiencies with low quality, nonuniform fuels; and (2) the pyrolytic burner, which has very low stack emissions with relatively high thermal efficiencies and wood combustion rates.

Present day wood-fired boiler systems are complicated and cost considerably more than a comparable petroleum-fired installation. As much as 25 percent of the capital cost is in the fuel-handling equipment, and another 20 percent is in the air pollution control system. Because of the high capital costs and the lower thermal efficiencies of burning wood compared to oil (68% vs. 82%), the success of the wood-fired systems depends on the low cost of the wood fuel supplies.

The North American forest products industry has been successful in developing useful technologies for recovering energy from wood residues as process heat or steam, which are cost-effective. New technologies promise even greater economic benefits.

Cogeneration Technology. Cogeneration is the concurrent generation of electricity and the use of exhaust heat, usually in the form of process steam, for manufacturing operations. This is done by burning fuel (in this case, wood) to make high-pressure steam, 600 to 1200 lb/in.^{2}, passing this steam through a back-pressure or extraction turbine to drive a generator, and then using the steam exhausted from the turbine at lower pressures, 50 to 300 lb/in.^{2}, for process heat. This technology gets full use of the energy contained in the fuel. Wood at 55 percent moisture will generate power at about 60 percent efficiency.

The forest products industry is a major user of cogeneration technology, as it requires

large quantities of process steam as well as electricity. It produces about 50 percent of its electricity needs in this way. The pulp and paper industry alone is the largest producer of energy by cogeneration of any U.S. industry. Electricity self-sufficiency is likely to increase to 80 to 90 percent for forest industries in the future through cogeneration. However, there is a minimum plant size for economical power generation. Steam usage should be more than 70,000 to 120,000 lb/hr, equivalent to 3 to 5 MW of back pressure, for a plant economically to employ cogeneration systems.

Saccharification–Fermentation

The saccharification–fermentation (SF) method for derivation of fuels and chemicals from wood is based on the breakdown or hydrolysis of the polysaccharides in wood to the constituent monomeric sugars. The six-carbon or hexose sugars (glucose, galactose, and mannose) then are fermentable to ethyl alcohol (ethanol or grain alcohol, C_2H_5OH) by yeast fermentation in the same way that ethanol is produced from grains or fruits. Obviously the concept is not a new one; the polysaccharide character of wood has been known for over 100 years. The limitations on the use of wood for ethanol production primarily have been the difficulties in separating and hydrolyzing the crystalline cellulose component in wood. Both acids and enzymes can be used to hydrolyze the cellulose to glucose, but only acids have been utilized commercially for wood hydrolysis to sugars, and only in foreign countries.[33]

In contrast to the Western economy, the Soviet Union continued expansion of wood hydrolysis facilities, and about 40 such plants are presently in operation. All the Russian plants are based on dilute sulfuric acid in percolation towers. In the West interest in producing alcohols from wood was revitalized by the dramatic increase in the price of petroleum in the 1970s and the push to decrease oil imports by substituting gasohol, which is one part alcohol in nine parts gasoline, for 100 percent gasoline at gas pumps. Both ethanol and methanol can be used in gasohol blends.

Because of the high oil prices, the country of Brazil (most of whose sugars are produced from sugarcane) took the dramatic step of shifting to a much greater use of fuel alcohol. One wood hydrolysis plant was constructed, but it was uneconomical to operate and was shut down. However, Brazilian experience has demonstrated that fermentation ethanol (95 percent ethanol and 5 percent water) is a perfectly satisfactory motor fuel. At least 500,000 Brazilian automobiles operate on undried alcohol continuously, and most of the rest of their fleet operates on this fuel on weekends when only alcohol is available at the gas stations.[33] A number of methods can be used for production of ethanol from wood, as described below.

Cellulose, the major component of wood, gives about a 90 percent yield of pure glucose under laboratory conditions of hydrolysis, according to the following equation:

$$(C_6H_{10}O_5)_n + nH_2O \xrightarrow{\text{acid}} nC_6H_{12}O_6$$

where n is in the range of 10,000 to 15,000. The hemicellulose fraction gives a mixture of sugars—xylose, arabinose, mannose, galactose, and glucose. Glucose, galactose, and mannose are yeast-fermentable sugars, whereas the pentoses (xylose and arabinose) are non-yeast-fermentable. The potential total reducing sugar yield from wood averages 65 to 70 percent, whereas the fermentable sugar yield is about 50 percent for the hardwoods and 58 percent for the softwoods. The lower quantity of fermentable sugar from the hardwoods is due to their higher content of pentosans, compared to the coniferous woods.

Hemicelluloses hydrolyze much more easily and rapidly than cellulose. Temperatures and acid concentrations that hydrolyze the cellulose to glucose in a matter of a few hours readily convert much of the hemicellulose into simple sugars in minutes or even seconds. Under industrial conditions of hydrolysis, the sugars formed undergo decomposition, with the pentoses decomposing more rapidly than the hexoses. Thus, the conditions of hydrolysis

cause variations in the ratio and yields of the various sugars due to (1) their different rates of formation by hydrolysis and (2) their different rates of decomposition.

The polysaccharides of wood (holocellulose) may be hydrolyzed by two general methods: (1) by strong acids, such as 70 to 72 percent sulfuric acid or 40 to 45 percent hydrochloric acid; or (2) by dilute acids, such as 0.5 to 2.0 percent sulfuric acid. The hydrolysis by strong acids is constant, proceeds as a first-order reaction, and is independent of the degree of polymerization. The reaction may be represented as follows:

$$\text{Holocellulose} \xrightarrow[\text{acid}]{\text{strong}} \left.\begin{matrix}\text{Swollen cellulose}\\ \text{Soluble pentosans}\end{matrix}\right\} \longrightarrow \text{Soluble polysaccharides}$$

$$\text{Soluble polysaccharides} \xrightarrow{\text{dilute acid}} \text{Simple sugars}$$

In dilute-acid hydrolysis, the reactions are heterogeneous and more complex than this because no swelling and solubilizing of the cellulose occurs. Cleavage of the insoluble cellulose results directly in low-molecular-weight oligosaccharides (intermediate products), which are rapidly converted to simple sugars, as indicated below:

$$\text{Holocellulose} \xrightarrow[\text{dilute acid}]{(1)} \begin{matrix}\text{Insoluble "stable" cellulose}\\ \text{Soluble hemicellulose intermediates}\\ \text{Pentose sugars}\end{matrix}$$

$$\begin{matrix}\text{Insoluble "stable" cellulose}\\ \text{Soluble hemicellulose intermediates}\\ \text{Pentose sugars}\end{matrix} \xrightarrow{(2)} \text{Oligosaccharides (cellulose intermediates)} \xrightarrow{(3)} \begin{matrix}\text{Hexose sugars}\\ \text{Pentose sugars}\end{matrix}$$

Reaction (1) is rapid and occurs under mild conditions, hydrolyzing mainly the hemicelluloses. Reaction (2) is slow, proceeds as a first-order reaction, and is the limiting reaction in this process. Reaction (3) is rapid.

Based on the above methods, two classes of industrial processes have been developed, namely, the Bergius-Rheinau process, based on the use of concentrated hydrochloric acid at ordinary temperatures, and the Scholler-Tornesch process, in which very dilute sulfuric acid is used at temperatures of 170 to 180°C (338–356°F). The latter method in an improved form is known as the Madison process, based on work done at the United States Forest Products Laboratory in Madison, Wisconsin. A number of modifications have been developed, including four in Japan.

Bergius-Rheinau Process. Hydrochloric acid of about 40 to 45 percent (by weight) is produced by reinforcing recovered, weaker acid with hydrogen chloride from salt-sulfuric acid reactors, or by burning chlorine with illuminating gas.

Wood chips are air-dried, then charged into a tile-lined reactor, and extracted countercurrently by the acid. The fresh strong acid enters that part of the battery of diffusers or reactors which contains the most nearly exhausted wood, and is pumped through the following containers until it is nearly saturated with the carbohydrates dissolved from the wood. Part of this solution is mixed, under slight cooling, with fresh wood, and the mixture is charged into the head container of the battery. After the container is filled and a few hours are allowed for reaction time, an amount of solution is forced from this container, through the pressure of the incoming acid, that corresponds to the yield from one charge.

The drawn-off solution contains hydrochloric acid and carbohydrates in about equal parts, at a concentration of about 25 percent (by weight) each. It then is concentrated in stoneware tubes under vacuum. The distillate, containing about 80 percent hydrochloric acid with minor proportions of acetic acid and furfural, is reused after fortification. The concentrated sugar solution is dried to a powder in a spray-dryer, where it also loses most of the remaining acid.

The dry, somewhat acid, powder contains the carbohydrates in the form of intermediate

polymers (oligosaccharides), which are water-soluble and must undergo further hydrolysis to yield simple sugars, either for fermentation or for crystallization. This is done by dissolving the oligosaccharides, diluting the solution to approximately 20 percent sugar concentration, and heating it for 2 hours in the presence of 2 percent acid at 125°C. Part of the glucose can be crystallized from the neutralized and reconcentrated solution, while the mother liquors are fermented to alcohol or used for growing yeast. A diagrammatic flow chart for the Bergius process is shown in Fig. 7.23.

When hardwoods are to be used, it is necessary to remove a part of the hemicellulose first by prehydrolysis. This has been done on a large scale with straw, a substance chemically similar to hardwood, by heating it in an 8-to-1 liquid-to-solid ratio, with 0.5 percent sulfuric acid for 2 to 3 hours at 130°C (266°F). Without prehydrolysis, hardwoods and straws form slimy materials, probably because of their high hemicellulose content, which prevent the flow of the hydrolyzing acid.

In the Bergius-Rheinau process the concentrated hydrochloric acid employed requires dried wood, and recovery of the acid is essential. The process gives high yields of sugars (to 65%) at high concentrations. The intermediate sugars first obtained, however, call for an extra processing step to reduce them to monomers, before fermentation or crystallization.

Two Japanese processes have been developed using concentrated (above 62%) sulfuric acid. In the Nihon-Mokuzai-Kagaku process, the acid is neutralized by lime, and the gypsum may be used for making gypsum board and for other purposes. In the Hokkaido process, the acid is recovered by dialysis. Two other Japanese developments are the Udic-Rheinau process, using 41 percent hydrochloric acid, and the Noguchi-Chisso process, using anhydrous hydrochloric acid. All four of these Japanese processes operate at atmospheric pressure and at temperatures below 100°C. All involve a prehydrolysis step to remove the hemicellulose because the main product is crystalline glucose.

Madison Process. In the Madison process, dilute sulfuric acid with an average concentration of 0.5 percent is pressed through wood in the form of sawdust and shavings. Regular flow of the acid and of the resulting sugar solution is one of the two principal requirements; the other is a lignin residue that can be discharged from the pressure vessels without manual labor. Both depend upon careful charging of the wood, which should not contain too many very fine particles, and upon maintaining a pressure differential of not more than 5 to 6 lb/in.^{2} between the top and the bottom of the digester.

The digesters or percolators are pressure hydrolyzing vessels, commonly employing a pressure of 150 lb/in.^{2}, and having a capacity of 2000 ft^3 each. In the original Scholler plants in Germany, the digesters were lined with lead and acid-proof brick; in the Madison process

Fig. 7.23. Diagrammatic flow sheet for the Bergius wood hydrolysis process.

a lining of "Everdur" metal was found to give sufficient protection.

The wood, about 15 tons, is pressed down with steam and then is heated by direct steam, after which the acid is introduced. The practice in Germany was to bring the dilute acid into the digester in several batches, with rest periods of about 30 minutes, heating the wood to temperatures of 130°C (266°F) at first, then to 180°C (356°F), while keeping the temperature of the entering acid 10 to 20°C lower. A total of about 14 hours was required to exhaust the wood, yielding about 50 lb of carbohydrates for 100 lb of dry wood substance. In the Madison process, continuous flow of the acid, and, correspondingly, of the sugar solution, is provided—in other words, continuous percolation. The cycle is thereby reduced to 6 hours, and the yields are increased somewhat. The lignin is blown out of the digester by opening the specially constructed bottom valves while the vessel is still under pressure.

The sugar solutions usually contain about 5 percent sulfuric acid. The solutions, still under pressure (150 lb/in.^{2}), are flash-evaporated to 35 lb/in.^{2}, neutralized with lime at that pressure, and filtered. Calcium sulfate is much less soluble at the elevated temperature corresponding to the pressure than it is at 100°C. This is a fortunate circumstance, for it must be removed to an extent sufficient to avoid difficulties caused by the formation of incrustations in the subsequent alcohol distillation. The filtered solution is cooled by further flash evaporation and heat-exchanged with water to the fermentation temperature.

Sugar yields from coniferous woods (softwoods) are about 50 percent at an average concentration of 5 percent. When fermented, the average ethyl alcohol yield per ton of drywoods is 50 to 60 gallons and sometimes higher.

The dilute sulfuric acid employed in the Madison process gives lower yield (49–55%) of sugars than the Bergius-Rheinau process, and only very dilute solutions are obtained directly. Recovery of heat is easier in the Madison process, and the acid need not be recovered.

The dilute-acid hydrolysis method presently is preferred for production on a commercial scale although the yield of glucose is only about 50 percent based on the cellulose. Nearly quantitative yields can be obtained by strong-acid hydrolysis, but are offset by higher capital costs for corrosion-resistant equipment and higher operating costs for acid recovery plus acid losses. More recent efforts to hydrolyze cellulosics with dilute acid have involved innovative reactor schemes. Rugg and Brenner described equipment for a screw-fed, continuous, single-stage hydrolyzer, and Thompson and Grethlein investigated a plug-flow reactor.[33] Both processes take advantage of the improved kinetics of cellulose hydrolysis relative to glucose decomposition at higher temperatures and shorter retention times to improve sugar yields. Both also have the advantage of being simple one-stage processes.

The U.S. Forest Products Laboratory (FPL) in cooperation with the Tennessee Valley Authority (TVA) has been studying a two-stage dilute acid hydrolysis process based in part on studies of Cederquist in Sweden during the 1950s. The first stage (prehydrolysis) selectively removes the hemicellulosic sugars with dilute sulfuric acid at about 170°C prior to hydrolysis of the lignocellulosic residue to glucose in the higher-temperature (230°C) second stage. The two-stage dilute acid process has a number of important advantages: (a) the carbohydrates are fractionated into hemicellulosic sugars and glucose, so that the separate utilization of each fraction is facilitated; (b) glucose is isolated in moderately good yield (~50%); (c) the glucose solution from the second stage is moderately concentrated (~10–12%); and (d) the consumption of acid and steam is relatively low.

Despite recent efforts such as these to produce glucose from agricultural residues and wood, little attention has been given to improving the kinetic modeling of cellulose saccharification with dilute acid. Kinetic modeling plays an important role in the design, development, and operation of processes to hydrolyze cellulosics to fermentable sugar.

Fig. 7.24. Model for dilute acid hydrolysis of cellulosics. (*From Connor, A. H., et al.,* in Cellulose: Structure, Modification and Hydrolysis, *R. A. Young and R. M. Rowell, Eds., copyright © John Wiley & Sons, New York, 1986; with permission of the publisher.*)

A new model for the dilute acid hydrolysis of cellulose was developed at FPL in connection with studies of the two-stage dilute sulfuric acid hydrolysis process (Fig. 7.24).[34] The model incorporates the effect of the neutralizing capacity of the substrate, the presence of readily hydrolyzable cellulose, and the reversion reactions of glucose in acid solution. Although general in nature, the model was developed specifically for application to the dilute, sulfuric acid hydrolysis of prehydrolyzed wood. A computer program simulating the new model can be used to predict yields of free glucose, reducing sugars, reversion material, remaining cellulose, and glucose loss due to dehydration as a function of acid concentration, temperature, and reaction time.

Enzymatic Hydrolysis. Saccharification of wood polysaccharides to sugars can be accomplished by enzymatic techniques instead of acid hydrolysis. The U.S. Army Natick Laboratories developed a method for conversion of cellulose to glucose with a cellulose enzyme from an active strain of the fungus *Trichoderma viride.* However, extensive pretreatment of wood is necessary before sufficient enzymatic hydrolysis will take place.

Enzymatic methods show the biggest promise for conversion of waste paper from municipal waste into glucose for ethanol production. Because paper is composed primarily of wood cellulose fibers, the enzyme inhibition due to lack of accessibility with whole wood is partially alleviated. As mentioned previously, waste paper can represent up to 50 percent of typical municipal waste. Currently the separated paper from the waste is burned for its fuel value.

The Gulf Oil Company developed a method called simultaneous saccharification and fermentation (SSF) for enzymatic conversion of waste paper to ethanol.[35] In this process the cellulose is enzymatically hydrolyzed and the glucose yeast-fermented in one operation. This modification, along with improved enzyme production and performance, has made the enzymatic technique more economically viable for conversion of waste paper to ethanol. The process was donated to the University of Arkansas for further development.

Fermentation of Sulfite Waste Liquor. The sulfurous acid used in the sulfite pulping liquor causes hydrolysis of the more easily hydrolyzable components of wood, especially the pentosans in the hemicellulose. About 35 percent of the potentially fermentable sugars in the wood are hydrolyzed. However, most of them are decomposed during the long pulping procedure so that only one-fourth to one-third of the sugars, including the more resistant hexoses, remain in the waste sulfite liquor. If these sugars are fermented by yeast,

about 12.5 gallons of 95 percent alcohol per ton of wood may be produced.

A large number of plants in Europe and a few in North America have been constructed to utilize the sugar hydrolyzate in sulfite waste liquors. The procedure as carried out in one American operation is as follows:

The liquor is recovered from the digester by discharging it in such a manner that as much sugar as possible is removed with a minimum of dilution by washing. Free sulfur dioxide is removed and recovered by blowing steam through the solution, which decreases the acidity from a pH of 2.2 to a pH of about 3.9. The liquors are cooled by a vacuum flash and neutralized by lime to a pH of about 4.2. A small amount of inorganic nitrogen is added for yeast growth, and about one percent by volume of yeast is added continuously. Fermentation is carried out in a series of tanks, the solution flowing from one to the other with agitation to keep the yeast in suspension. The yeast is recovered by centrifuging and mixed with the new sugar solution entering the fermenter. About 30 hours is required for fermentation. The alcohol content of the fermented liquor is about one percent by volume. The ethyl alcohol is recovered in stainless steel stills. Methanol and other alcohols are obtained in small amounts as by-products.

Thermal Decomposition

When wood is heated in the absence of air or with only limited amounts of air, thermal degradation takes place. This begins at about 100°C and increases with rising temperature. At about 270°C exothermic reactions set in, causing a rise in temperature (usually held at 400–500°C) bringing about complete carbonization. The products are charcoal, condensable liquids, and noncondensable gases.[36] The condensable liquids separate into aqueous (pyroligneous acid) and oil and tar fractions (Fig. 7.25). The charcoal, the gases (low Btu gas), and the oil fractions all may be used as fuels.

A number of terms are used for the thermal decomposition of wood and generally refer to similar processing methods: *carbonization*, *pyrolysis*, *gasification*, *wood distillation*, *destructive distillation*, and *dry distillation*. All result in the thermal breakdown of the wood polymers to smaller molecules in quantities dependent on reaction conditions. The hydrogen content of the gas increases with increasing temperature of pyrolysis. The wood gas has a fuel value of 300 Btu/ft^3.

The yields of the different materials obtained by pyrolysis vary with the species of woods used and the type of equipment and system employed. Manipulation of three variables, mainly the temperature, heating rate, and gas residence time, can greatly alter the relative proportions of the gas, liquid, and char produced. For example, the oil and tar yields can be varied between 1 and 40 percent and the char between 40 and 10 percent or less. Low temperatures favor liquids and char, low heating rates favor gas and char, and short gas residence favors liquids. Conversely, high temperatures favor gas, high heating rates favor liquids, and long gas residence times favor gas. Thus, the various product fractions can be preferentially manipulated by proper combinations of these variables.

During World War II in Germany, automobiles were fueled by the gases produced from thermal decomposition of wood; and research is ongoing today on the more efficient gasification of wood. Destructive distillation has been used throughout most recorded history to obtain turpentine from pinewood, as discussed later in the chapter.

Low Btu Gas and Oil. There has been considerable research activity to produce low Btu gas and oil from wood for energy. The Georgia Institute of Technology and the Tech-Air Corporation have intensively studied pyrolysis of biomass materials and have built several pilot plants. Their pyrolytic process (GT/T-A) has been applied to forestry residues. Their plant differs from the older wood-distillation plants, such as the Badger-Stafford retorts used in the 1930s, in several respects, such as the use of smaller-sized equipment to process 7 dry tons/hr compared to 3 to 4 tons/hr, the use of wood dried to 7% moisture

Fig. 7.25. Products obtained from the thermal decomposition of wood. (*From* Introduction to Forest Science, *2nd ed., R. A. Young and R. L. Giese, Eds., copyright © 1982 John Wiley & Sons, New York; with permission of the publisher.*)

instead of to less than 0.5%, and the use of a small quantity of air inside this retort (approx. 0.25 lb/lb feed) to sustain the reaction instead of using gases in order to heat it to 1000°F. The on-line time of the rated capacity for the GT/T-A system is claimed to be better than 90 percent, as compared to 67 percent for the Badger-Stafford units.

For the GT/T-A system the wood material is hogged so that the maximum size is not more than about one inch in any dimension. Sawdust may be used directly. The relatively uniform material is conveyed to a dryer where it is dried to about 7 percent water content. The dryer is heated by a portion of the wood gas from the reactor, and should this not be sufficient, the oil produced during pyrolysis can be used as a backup fuel.

The dried wood is fed into the top of the reactor though an air-lock and moves downward by gravity. A sensing device measures the bed height and controls the input of the wood. The temperature increases from 350 to 500°F at the top of the bed to 1000 to 1700°F in the pyrolysis zone. The gases move upward through the bed of descending wood and leave the top at 350 to 500°F. The reactor is operated at slightly under atmospheric pressure by an induced draft fan. The charcoal passes through an outfeed device at the bottom of the reactor into a sealed chamber, where it is cooled by a water spray and then conveyed to a storage bin. The rate of charcoal discharge controls the rate of passage of material though the unit.

The gases from the reactor pass into a scrubber where they are sprayed with cooled pyrolytic oil, which removes particulates and cools the gas stream to 180 to 200°F, causing condensation of the condensable organic substances into an oil mixture. The oil is filtered, the filter cake returned to the reactor,

and the filtered oil pumped to a tank, cooled, and then recirculated through the scrubber. Excess oil is pumped to a storage tank.

The gases from the scrubber-condenser contain the noncondensable gases, low-boiling organic vapors, and water vapor. A portion of the gases is used to heat the wood dryer, and the remainder are available as a fuel for other purposes, but it is desirable that they be used near the pyrolysis plant.

Operating conditions can be controlled to vary the distribution of the energy among the products: gas, oil, and charcoal. Thus, high yields of charcoal can be obtained with correspondingly reduced yields of oil and gas and vice versa, or the unit may be operated as a gasifier with only a 3.8 percent yield of charcoal. The off-gases in the latter case can be burned directly in a gas-fired boiler, or the condensable oils can be removed first and used separately as fuel. The total energy recovery in the form of wood gas, oil, and charcoal is about 95 percent of the energy in the dry wood used in the process.

Gasification. Gasification is the thermal degradation of wood or other carbonaceous material in the presence of controlled amounts of oxidizing agents, such as air or pure oxygen. It is carried out at higher temperatures than those used in the pyrolysis process, up to around 1000°C. Hence, the reaction rates are very fast, making equipment design critical. The thermal efficiency for conversion of wood to gases is 60 to 80 percent, which compares favorably with that of coal. Wood gasification offers several advantages over coal: (1) much lower oxygen requirements, (2) practically no steam requirements, (3) lower costs for changing H_2/CO ratios, which are already higher in wood gas, and (4) no or very little desulfurization costs. Coal has an advantage in that larger plants can be built for coal gasification than is normally the case for wood, because of its procurement advantages.

As noted above, the GT/T-A pyrolysis system also can be used as a gasifier by increasing the amount of air to the reactor and screening the charcoal and recycling the coarser particles to the incoming feed material. In this way, most of the wood is converted to gas and oil. If all the off-gases from the converter are conducted directly into a boiler as a hot gaseous fuel, then the GT/T-A process is essentially a gasifier. This requires that the gases be utilized as a fuel close to the pyrolysis plant. The use of air in the reactor results in the presence of nitrogen in the gas, causing a reduction in fuel value from 360 to 420 Btu/ft^3 (without air) to 140 to 200 Btu/ft^3 (with air).

Several gasifiers have been designed to handle municipal refuse, wood wastes, and other biomass materials. Basically, gasifiers fall into two types: (1) those that use air and (2) those using oxygen. An example of the air system is the Moore-Canada gasifier. The wood residues are carried on a moving bed through stages of drying, reduction, and char oxidation, and the ash is discharged in granular form. The maximum temperature in the reaction zone is about 1222°C. The hydrogen content of the crude gas is increased from 8 to 10 percent up to 18 to 22 percent by adding steam to the air intake.

The Union Carbide Company Purox gasifier is an example of the oxygen system. This unit also uses a moving-bed reactor. Pure oxygen is the oxidizing agent used to convert the char into CO and CO_2. Molten ash leaves the bottom at about 1670°C.

If air is used, the crude gases contain about 46 percent nitrogen, which must be removed by cryogenic means; but if oxygen is used, it first must be separated from air, with oxygen and nitrogen separated by a cryogenic system. (Chapter 13). The crude gases also contain an oil and tar fraction, about 2 percent of the wood (dry basis). Gasifiers designed for wood operate at atmospheric pressure in contrast to coal gasifiers, which operate at pressures up to 400 psi.

The composition of the wood gas varies according to the technology used. If the limited oxygen required is supplied by air, a typical gas might contain 10 to 18 percent H_2, 22 to 30 percent CO, 6 to 9 percent CO_2, 45 to 50 percent N_2, and 3 to 5 percent hydrocarbons and have a heating value of about 1700 kcal/m^3 (180 Btu/ft^3). If pure oxygen is supplied to the process, the gas

might contain 24 to 26 percent H_2, 40 percent CO, 10 percent hydrocarbons, and 23 to 25 percent CO_2 and have a heating value of 2900 kcal/m^3 (350 Btu/ft^3). The gas may be used directly as a fuel in a variety of ways; for example, supplying boiler energy in an industrial plant at the rate of 2.5×10^8 kcal/hr from about 136 kg/day of dry wood.

The technology of gasification is under active development in both equipment and process design to reduce capital costs, and in the chemistry of the process to improve yields and obtain more favorable gas ratios. The latter includes the use of catalysts to enhance the process and to promote the formation of specific products, such as methane or ethylene for increased thermal efficiency or for feedstocks for chemical synthesis.

Methanol. As is the case with ethanol, the concept of producing methanol from wood is now new. Methanol obtained from the destructive distillation of wood represented the only commercial source until the 1920s. The yield of methanol from wood by this method is low, only about 1 to 2 percent or 20 liters/metric ton (6 gal/ton) for hardwoods and about one-half that for softwoods. With the introduction of natural gas technology, the industry gradually switched to a synthetic methanol formed from a synthesis gas (syngas) produced from reformed natural gas. Two volumes of H_2 and one volume of CO are reacted in a catalytic converter at pressures of 1500 to 4000 psi to produce methanol. Presently, 99 percent of the methanol produced in the United States is derived from natural gas or petroleum.

Methanol also may be produced from wood gas; so wood could be a future raw material for making methanol, especially for use as an additive to gasoline for internal combustion engines. Thus reforming the gasification products obtained at high temperatures is a second method for the production of methanol from wood. This is in contrast to the older method (destructive distillation), which directly yields small quantities of methanol at lower temperatures as mentioned above.

If wood gas is produced in a gasifier using air as the oxidizing agent, the nitrogen (about 40%), as well as the CO_2 must be removed. The CO_2 is removed by passing the gas through hot potassium carbonate followed by scrubbing with monoethanolamine. The nitrogen, hydrocarbons, residual CO_2, and water are removed by freezing them out in a cryogenic system. By means of an iron catalyst, a "shift conversion" of a portion of the CO is carried out with steam in order to reach the required two parts of H_2 to one part of CO, as follows:

$$CO + H_2O \xrightarrow{\text{catal.}} H_2 + CO_2$$

$$\Delta H = -590 \text{ Btu/lb CO}$$

The pressurized gas then is led to the methanol reactor. Two different catalyst systems may be used: (1) a zinc–chromium catalyst requiring gas pressures of 2000 to 4000 psi or (2) a copper catalyst system at 1000 to 2000 psi. About 95 percent of the gas is converted to methanol by this exothermic reaction:

$$2H_2 + CO \xrightarrow{\text{catal.}} CH_3OH$$

$$\Delta H = -1200 \text{ Btu/lb CO}$$

The yield of methanol with present technology is about 390 liters/ton of oven dry wood. Conversion of wood to methanol is less efficient than producing methanol from natural gas or coal—about 38 percent compared to about 60 percent, based on the heat value of methanol as a percent of the total energy input into the plant.

The raw materials cost is the most significant operating cost. The production cost for methanol from wood is estimated to be almost twice that produced from natural gas. Improving the efficiency of the gasifier reactor to increase the quantities of CO and H_2 produced from wood would enhance the process.

Charcoal and Other Chemicals. Production of charcoal and tars by destructive distillation is the oldest of all chemical wood-processing methods. Charcoal probably was first discovered when the black material left over from a previous fire burned with intense heat and

little smoke and flame. For centuries, charcoal has been used in braziers for heating purposes. Destructive distillation of hardwoods has been carried out with charcoal the product sought and volatiles as by-products; with softwoods (pines), volatiles were the principal products (naval stores), with charcoal considered a by-product.

In the United States, charcoal production began in early colonial days. During this period, principal uses of charcoal were as a fuel in blast furnaces for the production of pig iron and as an ingredient of gunpowder. Charcoal needed by the iron industry had to have a high crushing strength, and thus was made from dense hardwoods such as maple, birch, oak, and hickory. A softer charcoal was preferred for making gunpowder and was produced from willow and basswood. The first successful blast furnace was built in Saugus, Massachusetts, in 1645. Other furnaces were started in close proximity to iron ore deposits. These early furnaces were small, usually producing only 1 to 3 metric tons of pig iron per day. In the late 1880s, wood charcoal lost its metallurgical market to coke, which was better suited to the demands of the larger furnaces then being built.

In 1812, the additional collection, by condensation, of the volatile substances from hardwood carbonization began; the products then were charcoal, crude pyroligneous acid, and noncondensable gases. The pyroligneous acid was refined to produce methanol, acetate of lime, which in turn was used to make either acetic acid or acetone and tar. The noncondensable gases in a normal wood distillation consisted of about 50 percent carbon dioxide, 30 percent carbon monoxide, 10 percent methane, 3 percent heavier hydrocarbons, and 3 percent hydrogen.[5] Table 7.10 gives the yields of products from an industrial operation. The tars and noncondensable gases usually were used as fuel. From the late 1800s until the 1920s, destructive distillation of hardwoods was an important source of industrial acetic acid, methanol, and acetone. This market was lost when these materials were made synthetically from petroleum. In 1920, there were approximately 100 plants recovering these products from hardwood distillation; the last of these plants ceased operation in 1969. The reader interested in details of the wood distillation processes is referred to the 1962 edition of *Riegel's Industrial Chemistry*, Chapter 15.

TABLE 7.10 Yield of Products from Destructive Distillation of One Ton of Dry Hardwood

Product	*Quantity*
Charcoal	270 kg (600 lb)
Noncondensable gas[a]	150 m^3 (5000 ft^3)
Soluble tar	83.6 liters (22 gal)
Pitch	29.7 kg (66 lb)
Creosote oil	12.5 liters (3.3 gal)
Methanol	11.8 liters (3.1 gal)
Acetone	2.7 liters (0.7 gal)
Allyl alcohol	0.2 liters (0.5 gal)
Ketones	0.8 liters (0.2 gal)
Methyl acetate	3.8 liters (1.0 gal)
Acetic acid	45.5 kg (101 lb)

Source: U.S.D.A. Forest Service Data.
[a]Composition in order of highest to lowest: CO_2, CO, CH_4, H_2, others.

In the early 1900s, charcoal from by-product recovery plants usually was used for cooking and heating in low-income areas and was known as a "poor man's fuel." Beginning in about 1950, there was an upturn in demand for charcoal for recreation use. In this era of suburban living, the use of charcoal briquettes for cookouts represents a significant market. The charcoal briquette now can be considered a luxury fuel, as it is too expensive for heating.

Basic techniques for producing charcoal have not changed over the years although the equipment has. Charcoal is produced when wood is burned under conditions in which the supply of oxygen is severely limited. "Carbonization" is a term that aptly describes the thermal decomposition of wood for this application. Decomposition of carbon compounds takes place as the temperature rises, leading to a solid residue that is richer in carbon than the original material. Wood has a carbon content of about 50 percent, whereas charcoal of a quality suitable for general market acceptance will be analyzed as follows: fixed carbon 74 to 81 percent, volatiles 18 to

23 percent, moisture 2 to 4 percent, and ash 1 to 4 percent. Charcoal with a volatile content over 24 percent will cause smoking and is undesirable for recreational uses.

Earthen "pit kilns" originally were used to produce charcoal. A circular mound-shaped pile of wood (15 to 45 cords) was built up with an open core 30 to 60 centimeters (1 to 2 feet) in diameter to serve as a flue. The entire surface of the pile, except for the top flue opening and several small openings around the bottom periphery, then was covered with dirt or sod sufficiently thick to exclude air. The mound was allowed to "coal" for 20 to 30 days to give the final product.

In the second half of the nineteenth century, brick or masonry "beehive" kilns came into widespread use. The capacity of many of these kilns was from 50 to 90 cords, and operation was essentially the same as for the pit kilns. Many other types of kilns have been used from time to time. Small portable sheet-metal kilns of 1 to 2 cords have been widely used, as have rectangular masonry block kilns of various designs. These were predominantly used by farmers and small woodlot owners.

Large-scale production of charcoal was done by distilling the wood in steel buggies in long horizontal ovens. The buggies rode in steel rails that carried the cars in line from predryers to the ovens and then to coolers. Charcoal was produced by this method in a matter of 24 hours. The latest types of charcoal-producing equipment are designed for continuous operation and make use of residues instead of roundwood. An example is the Herreshoff multiple hearth furnace, in which several hearths or burning chambers are stacked on top of one another, the number depending on capacity. Production in this type of furnace is from 1 to $2\frac{1}{2}$ metric tons of charcoal per hour.

Thermochemical Liquefaction

Although a reasonable amount of research effort has been expended on thermochemical liquefaction of wood, extensive commercialization of this process is not anticipated in the near future. The basis of the method is a high-pressure and high-temperature treatment of wood chips in the presence of hydrogen gas or syngas to produce an oil instead of a gas. The low-grade oil produced potentially could be substituted for some present petroleum uses.

An oil of a heating value of about 35,200 Btu/kg can be obtained by reaction of wood waste for one hour with syngas, a catalyst, a temperature of 750°F, and a pressure of 5000 psi. The feasibility of the process has been tested in a pilot plant in Albany, Oregon, based on laboratory work conducted at the U.S. Bureau of Mines. So far, it has been determined that a barrel of oil equivalent to No. 6 bunker fuel can be produced from about 405 kg of wood chips.[37]

A similar process was developed in Japan in the 1950s specifically to degrade lignin. Called the Noguchi process, it was thought to hold promise for production of phenols from lignin. By 1955, the Japanese investigators had discovered superior catalysts that converted a substantial portion of the lignin into a relatively few phenols. The Crown-Zellerbach Corporation (now defunct) subsequently obtained an option on the process and initiated its own trials. Despite several improvements researchers were able to make, the process did not prove profitable at the time. However, the company was routinely able to obtain a yield of 55 percent (and up to 65%) of distillable products. The major drawback was the inability to separate cleanly even the few different phenols remaining after the reactions.

NAVAL STORES

The U.S. naval stores industry began in the very early colonial days, when wooden vessels used tar and pitch from the crude gum or oleoresin collected from the wounds of living pine trees. The demand for tar and pitch from crude gum is now of minor importance.

The industry is centered in the southeastern United States and is confined to the longleaf and slash pine areas. There is also a small, but locally important, naval stores–producing

area in the Landes region of southwestern France, based on the maritime pine.

There are three routes by which naval stores are produced. The oldest method is the tapping of living trees to cause a flow of oleoresin. The second method is removal of naval stores by solvent extraction. The latter process now has replaced steam distillation as a means of recovering turpentine. In the United States, the latest and now the most important route is kraft (sulphate) pulping of pine, during which turpentine and tall oil are recovered as by-products of kraft pulp manufacture. A fourth process, no longer used in the United States, is recovery of turpentine and pine oils by the destructive distillation of pine wood.[38]

Worldwide, about 60 percent of naval stores are produced by tapping living trees, whereas gum naval stores account for only 4 percent of U.S. production. The sulfate process is now the major U.S. process.

Turpentine is a volatile oil consisting primarily of terpene hydrocarbons, having the empirical formula $C_{10}H_{16}$. These 26 atoms can have many different arrangements, only six of which are present in appreciable amounts in commercial turpentines: alpha-pinene (b.p. 156°C), beta-pinene (b.p. 164°C), camphene (b.p. 159°C), Δ^3-carene (b.p. 170°C), dipentene (b.p. 176°C), and terpinoline (b.p. 188°C). The molecular configurations of some of these are shown in Fig. 7.26.

Gum and sulfate turpentines have similar compositions. Gum turpentine contains 60 to 65 percent α-pinene, 25 to 35 percent β-pinene, and 5 to 8 percent other terpenes, compared to 60 to 70 percent α-pinene, 20 to 25 percent

Fig. 7.26. Some reactions of alpha-pinene and beta-pinene. (From *Goldblatt, L. A., "Yearbook of Agriculture," U.S.D.A., 1950–51*).

β-pinene, and 6 to 12 percent other terpenes for sulfate turpentine. Wood turpentine, which has 75 to 80 percent α-pinene, has no or very little β-pinene (0–2%), and also contains 4 to 8 percent camphene and 15 to 20 percent other terpenes.

Rosin, the other major naval stores product, is a brittle solid that softens at 80°C. Chemically it is composed of about 90 percent resin acids and 10 percent neutral matter. The resin acids are mainly *l*-abietic acid and its isomers, $C_{20}H_{30}O_2$. These are tricyclic monocarboxylic acids and are diterpenes.

Rosin is graded and sold on the basis of color, the color grades ranging from pale yellow to dark red (almost black). The color is due almost entirely to iron contamination and oxidation products. Fresh oleoresin, as it exudes from the tree, will yield a rosin that is nearly colorless. Color-bodies are removed by selective solvents and selective absorption from a 10 to 15 percent gasoline solution passed through beds of diatomaceous earth. About 70 percent of the world's rosin is produced in the United States. Naval stores also include important fatty acids, as well as steroids and other products.

Gum Naval Stores

The crude gum or oleoresin is obtained from healthy trees by exposing the sapwood. The lower part of the tree is faced, that is, a section of bark is removed, giving a flat wood surface for the gutters, which are inserted into a slanting cut made by a special ax. The gutters conduct the gum to a container that can hold 1 to 2 quarts of gum. At the top of the exposed face, a new V-shaped strip of bark is removed about every two weeks.

The operations of inserting gutters, hanging cups, and cutting the first bark preferably are done in December or January, as early facing stimulates early season gum flow. The gum continues to flow until November, with the height of the season being from March to September. The collected gum is distilled from a copper still; turpentine and water pass over, and the rosin is left in the still. The remaining molten rosin, plus impurities, is passed through a series of strainers and cotton batting to remove dirt particles. Then the liquid rosin is run into tank cars, drums, or multiwall paper bags for shipment.

Increases in yield of naval stores are brought about by chemical treatment of the exposed wood, especially with paraquat herbicides (dipyridyl compounds). This treatment stimulates extensive oleoresin formation and diffusion into the wood, extending to the pith of the tree and several feet above the treatment level. As much as 40 percent oleoresin content in the wood has been produced. Such treatment could double naval stores production, for both gum and sulfate processes. It also has the potential of providing a new type of wood naval stores by solvent extraction prior to kraft pulping, or a combination of both methods.

Wood Naval Stores

"Wood" naval stores are produced by solvent extraction of resin-rich wood from old southern pine stumps and roots. The depletion of these stumps from the large trees of virgin forests, combined with high labor costs, has brought about a major decline in the production of wood naval stores by this process.

In modern practice, all the resin products are removed from the shredded wood by solvent extraction; the solvent retained by the extracted wood chips is recovered by steaming. Extraction is carried out with naphtha (b.p. 90–115°C fraction). Multiple extraction is carried out in a series of vertical extractors in a countercurrent manner, whereby fresh solvent is used for the final extraction of a charge.

The solution from the extractors is vacuum-distilled and the solvent recovered. The remaining terpene oils are fractionally redistilled under vacuum and recovered as turpentine, dipentene, and pine oil. The nonvolatile rosin is of dark color and is upgraded by clarification methods, such as selective absorption of its solution (bed-filtering).

Sulfate Naval Stores

Sulfate turpentine is obtained as a by-product during the kraft pulping of pine woods. Vapors periodically released from the top of the digesters are condensed, and the oily turpentine layer is separated and purified by fractional distillation and treatment with chemicals to remove traces of sulfur compounds. Sulfate turpentine is very similar to gum turpentine obtained from the oleoresin of the tree and contains 60 to 70 percent α-pinene, 20 to 25 percent β-pinene, and 6 to 12 percent other pinenes. Sulfate turpentine from western North America woods contains appreciable amounts of Δ^3-carene, which is used as a solvent. In the United States, about 80 percent of the annual production of turpentine is from the kraft pulping of southern pines (approximately 125 million liters).

The spent black liquor from the kraft pulping of pines contains the less volatile products of the wood resin in the form of sodium salts or soaps. The liquor first is concentrated in multiple-effect evaporators, and then the concentrate is sent to settling tanks. The soaps rise to the surface, are skimmed off, and then are acidified with sulfurous or sulfuric acid. The crude tall oil rises to the top and is mechanically separated. Crude tall oil from southern pines contains 40 to 60 percent resin acids and 40 to 55 percent fatty acids with 5 to 10 percent neutral substances. These components are separated by fractional distillation under vacuum.

One metric ton of crude tall oil yields about 350 kg of rosin, 300 kg of fatty acids, and 300 kg of head and pitch fractions. For each metric ton of pulp produced, northern pines yield about 50 kg of tall oil, and the southern pines yield about 125 kg. The U.S. capacity for fractional distillation of tall oil is nearly one million metric tons per year.

The rosin component of tall oil is mostly made up of resin acids, which are diterpene derivatives. The major compounds (figures indicate averages) are as follow:

1. Abietic type—abietic (32%), neoabietic (4%), palustric (10%), and dehydroabietic (30%) acids.
2. Pimaric type—pimaric (4%), isopimaric (10%), and sandara copimaric acids (small amounts).

The fatty acids from tall oil have the following components: oleic (50%), linoleic (35%), conjugated linoleic (8%), stearic (2%), palmitic (1%), and others (4%). From tall oil heads, a fraction is produced that is composed predominantly of saturated acids, containing 55 percent palmitic acid. Oleic acid is distilled in grades that are 99.5 percent pure.

The neutral or "unsaponifiable" materials present in tall oil include anhydrides, phenolics, diterpene aldehydes and alcohols, stilbenes, and steroids. In the neutral fraction of southern pine tall oil soap, 80 compounds have been identified. They include 25.1 percent sistosterol and a total of 32.4 percent steroids. The sistosterol content of crude tall oil is 2 to 3 percent and is the main component of the neutral fraction.

Uses of Naval Stores Products

Navel stores products have a wide range of uses from ordinary household commodities to complex industrial applications, as outlined below.[38]

Turpentine. Turpentine is used directly as a solvent, thinner, or additive for paints, varnishes, enamels, waxes, polishes, disinfectants, soaps, pharmaceuticals, wood stains, sealing wax, inks, and crayons, and as a general solvent. The chemistry of its monoterpenes offers many possibilities for conversion to other substances, as illustrated in Fig. 7.26. There is increasing use of turpentine to produce fine chemicals for flavors and fragrances. An important use of turpentine is in conversion by mineral acids to synthetic pine oil. It also is a raw material for making terpin hydrate, resins, camphene, insecticides, and other useful commodities. These uses are included in the following summary of its applications:

- Solvents for paints, etc. (11%).
- Synthetic pine oil (48%), used for mineral flotation, textile processing, solvents,

odorants, bactericides, and conversion to their chemicals such as terpin hydrate, etc.
- Polyterpene resins (600–1500 mol. wt.) (16%), used for adhesives, pressure-sensitive sizes (dry cleaning, paper, chewing gum).
- Camphene → toxophene insecticides (16%).
- Flavor and fragrance essential oils (9%).

Dipentene. Dipentene is present in the higher-boiling fractions of wood turpentine. It is used in paints and varnishes and as a penetrating and softening agent in rubber reclamation.

Pine Oil. Pine oil obtained from wood naval stores has similar uses to those of the synthetic pine oil made from turpentine, given above.

Rosin. Rosin is used mainly in some modified form. Because the abietic-type acids in rosin each contain a carboxyl group and double bonds, they are reactive and can be used to produce salts, soaps, esters, amines, amides, nitriles, and Diels-Alder adducts; and they can be isomerized, disproportionated, hydrogenated, dimerized, and polymerized. When destructively distilled, rosin produces a viscous liquid, termed rosin oil, used in lubricating greases.

The paper industry uses large amounts of the sodium salt of rosin as paper size, which accounts for the greatest single use of rosin. The synthetic rubber industry is the second most important user of rosin. In making styrene-butadiene rubber, disproportionated rosin soaps are used alone or in combination with fatty acid soaps as emulsifiers in the polymerization process. Disproportionation decreases the number of double bonds in the abietic acid of the rosin, making a more stable material.

The adhesives industry is the third most important market for rosin. Rosin, modified rosins, and rosin derivatives are used in several types of adhesives, including the pressure-sensitive, hot-melt, and elastomer-based latices, and solvent rubber cements.

Protective coatings are the fourth major user of rosin, either directly or in a modified or derivative form. Varnishes and alkyds are the most common types of protective coatings using rosin. Rosin is combined with a heat-reactive phenol-formaldehyde resin to produce a widely used varnish. Printing inks also use substantial amounts of rosin.

The above rosin uses are summarized as follows:

- Paper size (33%).
- Chemical intermediates and rubber (42%).
- Resins and ester gums (18%).
- Coatings (3.9%).
- Other uses (3.1%).

Fatty Acids. Of the total fatty acids produced annually in the United States, amounting to more than 450,000 tons, 35 percent come from tall oil. The solvent extraction of pine wood yields only one percent fatty acids and their esters. The yield is not increased, however, by paraquat (dipyridyl herbicides) treatment. Hence, the kraft-pulping industry will continue to be the major source of fatty acids from wood. The approximate distribution of uses of the fatty acids are:

- Intermediate chemicals (43%).
- Protective coatings (28%).
- Soaps and detergents (11%).
- Flotation (3%).
- Other uses (15%).

A future product from the neutral fraction of tall oil may be sistosterol. This chemical has potential use in the synthesis of cortisones and other steroids and hormones by fermentation processes.

ADDITIONAL CHEMICALS FROM WOOD

A variety of additional chemicals are derived directly from wood or as by-products from the pulp and paper industry.

Tannins and Other Extractives

The tissues of wood, bark, and the leaves of trees contain a great variety of chemical

substances of considerable scientific interest and some of practical value. Turpentine, pine oil, and rosin from the resins of pines are the most important commercial extractives from American woods.

Tannin is a commercially important substance that can be extracted from the wood, bark, or leaves of certain trees and other plants. Tannins are complex dark-colored polyhydroxy phenolic compounds, related to catechol or pyrogallol, and vary in composition from species to species. They have the important property of combining with the proteins of animal skins to produce leather.

For many years, most of the leather in the United States was tanned with domestic tannins from hemlock and oak bark and from chestnut wood. Today only a small amount of tannin comes from these and other domestic sources. The most important source of vegetable tannin today is the wood of the quebracho tree, which grows mainly in Paraguay and Argentina. The tannin content of this tree and a few other sources of vegetable tannin are shown in Table 7.11.

The wood or bark for tannin production is reduced to chips and shreds by passing the material through hoggers or hammer mills. Then it is extracted with warm water in diffusion batteries. The dilute solutions are evaporated to the desired concentration. Loss of solubility of the tannin can be counteracted by treatment of the concentrate with sodium sulfite.

TABLE 7.11 Tannin Content of Some Plant Materials

Plant Material	*Percent Tannin*
Domestic sources	
Eastern hemlock bark	9–13
Western hemlock bark	10–20
Tanbark oak	15–16
Chestnut oak	10–14
Black oak	8–12
Chestnut wood	4–15
Sumac leaves	25–32
Foreign sources	
Quebracho heartwood	20–30
Mangrove bark	15–42
Wattle (*acacia* bark)	15–50
Myrobalan nuts	30–40
Sicilian suma leaves	25–30

Furfural

An additional, potentially important chemical derivable from wood as a result of hydrolytic (acid) treatment is furfural. Furfural is derived from the hemicellulose fraction of wood—specifically from the five-carbon or pentose sugars (primarily xylose). The pentose sugars are not yeast-fermentable by standard methods to ethanol but can be treated with acid, which causes dehydration and yields furfural.

$$\text{Hemicellulose} \xrightarrow{\text{Acid}} \text{Pentosans} \xrightarrow{\text{Acid}} \text{Furfural} \xrightarrow{\text{Reduce}} \text{Furfuryl alcohol}$$

Most furfural is produced from corncobs and oat and rice hulls, primarily by the Quaker Oats Company. The product is used in the chemical industry as a solvent and in wood rosin refining. A large amount of furfural is treated further to give furfuryl alcohol. The furfuryl alcohol is added to urea-formaldehyde resins in applications for adhesives and foundry core binders.

Vanillin

Vanillin is not a product of hydrolysis or of fermentation; in fact, it does not originate from the holocellulose, but rather than the lignin portion of the wood. The major organic material in sulfite waste liquor (spent liquor, after pulping wood by the sulfite process) is the lignin dissolved from the wood as lignosulfonic acid. Alkaline degradation of this lignin product produces vanillin, the same substance that occurs naturally in the vanilla bean:

Vanillin

With the Howard process, 5 to 10 percent vanillin is produced, based on the lignin in

the waste sulfite liquor. Some vanillin is produced from sulfite waste liquor in both the United States and Canada.

Dimethyl Sulfide and DMSO

Dimethyl sulfide (DMS) is recovered from the waste black liquor from the kraft pulping process by flash drying of the liquor after the DMS content has been increased by additions of sulfur. DMS is an odorant for natural gas and a solvent. DMS is oxidized to give dimethylsulfoxide (DMSO), an important solvent and a medicinal of interest in the treatment of arthritis.

Medicinals

The bark of the cascara tree of the northwestern region of the United States yields cascara, a laxative used in medicine. Several hundred tons of bark are harvested annually. Recently, one of the most promising anticarcinogenic compounds to be discovered in the last 15 years was found in the bark and weed of the Pacific Yew tree from the Pacific Northwest of the United States. This compound, taxol, is now being evaluated on a large scale by the National Cancer Institute.

The red gum tree of the southern United States exudes a yellowish balsamic liquid or gum from wounds, which is known as storax. It is produced by removing a section of bark and incising the wood in much the same manner as that used for the production of naval stores gum described above. Storax is used in medicinal and pharmaceutical preparations, such as adhesives and salves, and as an incense, in perfuming powders and soaps, and for flavoring tobacco.

Biotechnology Chemicals

As with pulping and bleaching, biotechnology too could have a considerable impact on the production of chemicals from wood and other forms of plant biomass. The effects of biotechnology probably will be noticed first in areas of enzymatic hydrolysis of polysaccharides and fermentation technology.

It should be possible to improve the efficiency of the cellulose enzyme complex for hydrolyzing cellulose to glucose. The enzyme complex apparently contains decrystallizing and hydrolysis enzymes that work together to convert cellulose to glucose. Isolation of the specific enzymes and genetic engineering could provide a more efficient complex.

As discussed earlier in this chapter, enzymes are the basis for yeast conversion of hexose (six-carbon) sugars, such as glucose and mannose, to ethanol. These enzymes also could be genetically engineered to improve the efficiency of alcohol production; several biotechnology firms are exploring this possibility.

The yeast (*Saccharomyces cerevisiae*) enzymes are specific to six-carbon sugars, but wood and other forms of biomass also contain large quantities of pentose sugars, especially xylose in hardwoods. The pentoses are not fermentable to ethyl alcohol with conventional yeasts. However, researchers at the U.S.D.A. Forest Products Laboratory have discovered a xylose-fermenting yeast (*Candida tropicalis*). Thus it now is possible to convert all wood sugars to ethyl alcohol with a combination of yeasts. Isolation of the specific enzymes and genetic engineering of the enzymes could dramatically improve the efficiency of this conversion.

Many other chemicals can be obtained from both yeast and bacteria fermentation of sugars and pulp mill effluents. Potential fermentation products from wood hydrolysates include acetone, organic acids (acetic, butyric, lactic), glycerol, butanediol, and others.[39]

REFERENCES

1. Young, R. A., and Giese, R. L. (Eds.), *Introduction to Forest Science*, 2nd ed., John Wiley & Sons, New York, 1990.
2. Anon., *Wood Handbook: Wood as an Engineering Material*, U.S.D.A. Agr. Handbook 72, U.S.D.A. For. Serv., Forest Products Laboratory, U.S. Govt. Printing Office, Washington, DC, 1987.
3. Anon., *Wood-Colors and Kinds*, U.S.D.A. Agr. Handbook 101, U.S.D.A. For. Serv., Forest Products Laboratory, U.S. Govt. Printing Office, Washington, DC, 1956.
4. Kubler, H., "Properties and Utilization of Wood," in *Introduction to Forest Science*, 2nd ed., R. A. Young and R. L. Giese (Eds.), pp. 451–477, John Wiley & Sons, New York, 1990.
5. Young, R. A., McGovern, J. N., and Rowell, R. M., "Wood for Fiber, Energy and Chemicals," in *Introduction to Forest Science*, 2nd ed., R. A. Young and R. L. Giese (Eds.), pp. 478–507, John Wiley & Sons, New York, 1990.
6. Anon., *Pulp & Paper*, p. 9 (Aug. 1989).
7. Kleinert, T. N., and Tayenthal, U. S., U.S. Patent 1,856,567 (1931).
8. Klausmeier, W. H., "Economic Analysis of an Innovative Process for Cellulose Acetate Production," in *Cellulose: Structure, Modification and Hydrolysis*, R. A. Young and R. M. Rowell, (Eds.), pp. 187–221, Wiley-Interscience, New York, 1986.
9. Herdle, C. E., Pancoast, L. H., and MacClaren, R. H., *Tappi*, **47** (10); 617 (1964).
10. Young, R. A., *Tappi*, **72**, 195 (1989).
11. Young, R. A., University of Wisconsin-Madison, unpublished results.
12. Springer, E. L., and Zoch, L. C., *Tappi*, **54** (12), 2059 (1971).
13. Johansson, A., Sachetto, J. P., and Roman, A., "*A New Process for the Fractionation of Biomass*," Battelle Geneva Labs, Geneva, Switzerland, 1982.
14. Clemont, L. P., *Tappi*, **53** (12), 2243 (1970).
15. Deltaes, G. G., and Lang, C. J., *Tappi*, **57** (5), 127 (1974).
16. Kleinert, T. N., *Tappi*, **57** (8), 99 (1974).
17. Katzen, R., Fredrickson, R., and Brush, B. F., *Chem. Eng. Prog.*, **76**, 62 (1980).
18. Paszner, L., and Chang, P. C., International Symposium on Wood and Pulping Chemistry, Tsakuba Science City, Japan, 1983.
19. Edel, V. E., *Dtsch. Papier Wirtschaft*, **1**, 39, (1984).
20. Aziz, S., and Sarkanen, K. V., *Tappi*, **72** (3), 169–175 (1989).
21. McCready, M. (Ed.), "Allcell Pulping: It Works!," *Paper Age*, p. 14 (Mar. 1991).
22. McKinney, R. W. N., "Recycling Fibre—the State of the Art in Europe," in *World Pulp & Paper Technology 1991*, F. Roberts (Ed.), pp. 31–37, The Sterling Publishing Group, London, 1991.
23. Stier, J. C., and Buongiorno, J., "The Forest Products Economy," in *Introduction to Forest Science*, R. A. Young and R. L. Giese (Eds.), pp. 508–531, John Wiley & Sons, New York, 1990.
24. Meyer, J. A., "Wood-Polymer Materials," in *The Chemistry of Solid Wood*, R. M. Rowell (Ed.), pp. 257–290, Am. Chem. Soc., Washington, DC, 1984.
25. Rowell, R. M., "Penetration and Reactivity of Cell Wall Components," in *The Chemistry of Solid Wood*, R. M. Rowell (Ed.), pp. 175–210, 1984.
26. Rowell, R. M., Tillman, A. M., and Simonson, R., *J. Wood Chem. Technol.*, **6** (3), 427–448 (1986).
27. Shiraishi, N., Aoki, T., Norimoto, M., and Okumuso, M., "Thermoplasticization of Cellulose and Wood," in *Graft Copolymerization of Lignocullulosic Fibers*, D. N.-S. Hon (Ed.), pp. 321–348, Am. Chem. Soc., Washington, DC, 1932.
28. Matsuda, H., *Wood Sci. Technol.*, **21**, 75 (1987).
29. Rowell, R. M., and R. A. Young, "Recent Advances in Lignocellulosic—Derived Composites," in *Assessment of Biobased Materials*, H. L. Chum (Ed.), pp. 2.1–2.23, Solar Energy Res. Inst., Apt. No. SERI/TR-234-3610, Dept. of Energy, Golden, CO, 1989.
30. Rowell, R. M., "Chemical Modification of Wood," *Commonwealth Forestry Bureau*, Oxford, England, **6** (12), 363–382 (1982).
31. Rowell, R. M., "Nonconventional Wood Preservation Methods," in *Wood Technology: Chemical Aspects*, Am. Chem. Soc., Symp. Ser. No. 43, 47–56, Washington, DC, 1977.
32. LeVan, S. L., "The Chemistry of Fire Retardancy," in *The Chemistry of Solid Wood*, R. M. Rowell (Ed.), pp. 531–574, Am. Chem. Soc., Washington, DC, 1984.
33. Wayman, M., "Comparative Effectiveness of Various Acids for Hydrolysis of Cellulose," in *Cellulose: Structure, Modification and Hydrolysis*, R. A. Young and R. M. Rowell (Eds.), pp. 265–279, Wiley-Interscience, New York, 1986.
34. Connor, A. H., Wood, B. V., Hill, C. G., and Harris, V. F., "Kinetic Modeling of the Saccharification of Prehydrolysyzed Southern Red Oak," in *Cellulose: Structure, Modification and Hydrolysis*, R. A. Young and R. M. Rowell (Eds.), pp. 281–296, Wiley-Interscience, New York, 1986.
35. Evert, R. G. H., Katzen, R., Fredrickson, R. E., and Kapisch, K. F., *Chem. Eng. Prog.*, **76** (9), 47 (1980).

36. Shafizadeh, F., "The Chemistry of Pyrolysis and Combustion," in *The Chemistry of Solid Wood*, R. M. Rowell (Ed.), pp. 489–530, Am. Chem. Soc., Washington, DC, 1984.
37. Appell, H. R., "The Production of Oil from Wood Waste," in *Fuels from Waste*, Academic Press, New York, 1977.
38. Zinkel, D. F., *J. Appl. Polym. Sci.*, **28**, 309–327 (1975).
39. Hajny, G. J., "Biological Utilization of Wood for Production of Chemicals and Foodstuffs," U.S.D.A. For. Serv., Res. Rept. FPL 385, Forest Products Laboratory, Madison, WI, 1981.

SELECTED REFERENCES

Books

American Wood Preservers' Association, *American Wood Preservers' Standards*, Washington, DC, 1971.
Britt, K. W., *Handbook of Pulp and Paper Technology*, 2nd ed., Van Nostrand Reinhold, New York, 1970.
Browning, B. L., *Methods of Wood Chemistry*, Vols. I and II, John Wiley & Sons, New York, 1967.
Casey, J. P., *Pulp and Paper*, 2nd ed., Vols. I, II, and III, John Wiley & Sons, New York, 1980.
Fengel, D., and Wegener, G., *Wood: Chemistry, Ultrastructure, Reactions*, Walter de Gruyter Pub., New York, 1984.
Glasser, W. G., and Sarkanen, S. (Eds.), *Lignin: Properties and Materials*, Am. Chem. Soc. Symp. Ser. 397, Am. Chem. Soc., Washington, DC, 1989.
Goldstein, I. (Ed.), *Organic Chemicals from Biomass*, CRC Press, Boca Raton, FL, 1981.
Hillis, W. E. (Ed.), *Wood Extractives*, Academic Press, New York, 1962.
Hon, D. N.-S., and Shiraiski, N., *Wood and Cellulosic Chemistry*, Marcel Dekker, New York, 1991.
Hunt, G. M., and Garratt, G. A., *Wood Preservation*, McGraw-Hill, New York, 1953.
MacDonald, R. G. (Ed.), *Pulp and Paper Manufacture*, 2nd ed., Vols. I, II, and III, McGraw-Hill, New York, 1969.
Rowell, R. M. (Ed.), *The Chemistry of Solid Wood*, Am. Chem. Soc., Washington, DC, 1984.
Rydholm, S. A., *Pulping Processes*, John Wiley & Sons, New York, 1965.
Sarkanen, K. V., and Ludwig, C. H. (Eds.), *Lignins*, John Wiley & Sons, New York, 1971.
Sjöström, E., *Wood Chemistry: Fundamentals and Applications*, Academic Press, New York, 1981.
Wenzel, H. F. J., *The Chemical Technology Wood*, Academic Press, New York, 1970.
Young, R. A., and Giese, R. L. (Eds.), *Introduction to Forest Science*, 2nd ed., John Wiley & Sons, New York, 1990.
Young, R. A., and Rowell, R. M., *Cellulose: Structure, Modification and Hydrolysis*, Wiley-Interscience, New York, 1986.

Articles and Chapters

Chum, H. L., "Structural Materials for the Automobile of the Future—Composite Materials," in *Assessment of Biobased Materials*, H. L. Chum (Ed.), pp. 1.1–1.19, Solar Energy Res. Inst., Rpt. No. SERI/TR-234-3650, Dept. of Energy, Golden, CO, 1989.
Jahn, E. C., and Stannett, V., "Polymer Modified Papers," in *Modern Materials*, Vol. II, H. H. Hausner (Ed.), Academic Press, New York, 1960.
Johansson, A., Aaltonen, O., and Ylinen, P., "Organosolv Pulping-Methods and Pulp Properties," *Biomass*, **13**, 45–65 (1987).
Levan, S. L., Ross, R. J., and Winandy, J. E., "Effects of Fire Retardant Chemicals on the Bending Properties of Wood at Elevated Temperatures," Res. Rept. FRL-RP-498, U.S.D.A., For. Serv., Forest Products Laboratory, 1990.
Shiraishi, N., Aoki, T., Norimoto, M., and Okumura, M., "Make Cellulosics Thermoplastic," *Chemtech.*, pp. 366–373 (June 1983).
Young, R. A., "Structure, Swelling and Bonding of Cellulose Fibers," in *Cellulose: Structure, Modification and Hydrolysis*, R. A. Young and R. M. Rowell (Eds.), pp. 91–128, Wiley-Interscience, New York, 1986.

8

Animal and Vegetable Fats, Oils and Waxes

Edmund W. Lusas and Khee Choon Rhee*

INTRODUCTION

Fats and oils predominantly are triesters (triacylglycerols, triglycerides) of glycerol and aliphatic fatty acids generally containing up to 24 carbon atoms. Waxes are esters of long-chain fatty acids, usually containing 24 to 28 carbon atoms, with long-chain primary alcohols (16 to 36 carbon atoms), or with alcohols of the steroid group.[1]

These compounds are members of a broader group of chemical substances called *lipids*, which has been classified by the National Research Council into: (1) *nonpolar lipids*—including esters of fatty acids (triacylglycerols and cholesteryl esters) that are virtually insoluble in water but soluble in most organic solvents, and enter metabolic pathways only after hydrolysis; and (2) *polar* or *amphipathic lipids*—including fatty acids, cholesterol, sphingolipids, and glycerophospholipids (mainly lecithins). The term *phospholipids* includes lecithins and sphingomyelins.[2] Other minor natural compounds, also extracted simultaneously by low-polarity organic solvents, include fat-soluble vitamins, colors, and flavors.

Fats and oils have a major role in human nutrition. They are concentrated dietary sources of energy, providing approximately 9 kcal/g when metabolized, compared to 4 kcal/g for carbohydrates and proteins, and account for about 36 percent of caloric intake per capita domestically.[1] Dietary lipids also can provide essential molecular structures that are synthesized by the body into compounds required for selective functioning of cell membranes and for regulation of life processes.

In the preparation of foods, fats and oils modify product texture, carry flavors, improve mouthfeel, provide a sensation of product richness, and induce satiety. They are used in many other commercial applications, including soaps and detergents, plastics and protective coatings, printing inks, and feeds for domesticated animals, and as carriers of pesticides for aerial spraying, for control of grain dust, and as feedstocks for chemical manufacturing processes.

*Food Protein Research and Development Center, Texas A&M University System, College Station, TX.

NOMENCLATURE, MOLECULAR STRUCTURES, AND POLYMORPHISM

General

The choice of the term "oil" or "fat" usually is based on the physical state of the material and tradition. Generally, oils are liquid at ambient temperatures, and fats are semisolid. Fats usually are of animal origin (beef tallow, pork lard and butter fat), while oils are extracted from fish or plant tissue or seeds. Nutritionists usually use the term "fat."

Over 95 percent of the weight of most extracted (*visible*) food fats are triacylglycerols formed by combination of glycerol (a trihydric alcohol) with three fatty acids, with loss of one molecule of water for each ester linkage:

$$\begin{array}{l} CH_2OH \\ | \\ CHOH + 3C_{17}H_{35}COOH = \\ | \\ CH_2OH \end{array}$$

Glycerol Stearic acids

$$\begin{array}{l} CH_2OOCC_{17}H_{35} \\ | \\ CHOOCC_{17}H_{35} + 3H_2O \\ | \\ CH_2OOCC_{17}H_{35} \end{array}$$

Glycerol tristearate
(tristearin)

Fatty Acids

The fatty acids are the building blocks of the triacylglycerols. Trace amounts of odd-number carbon fatty acids are found in most fats, and these compounds also have been synthesized for research purposes. Microorganisms frequently produce branched fatty acids. Heptadecenoic (17 carbon) acid is a major component of *Candida tropicalis* yeast fat. More than 90 percent of the fatty acids have an even number of carbon atoms, and are in aliphatic chains ranging from 4 to 24 carbons in length.

The names of common fatty acids under several conventions, carbon numbers, and selected properties are shown in Table 8.1. The common (trivial) names of some fatty acids are of long standing, and often indicate the initial source studied. As examples: butyric acid is a major component of butter flavor; the 6, 8, and 10 saturated fatty acids have been called the *goaty acids* because they impact characteristic flavors of goat and sheep milk and cheese; and the terms *olein* and *stearin* were applied in the early manufacture of oleomargarine and compounded shortenings to the liquid and solid fractions, respectively, of tallow separated by pressing.

Fatty acids sometimes are designated by the number of carbon atoms in the chain, followed by a colon and additional numbers indicating the number of double bonds. Thus, in the 18-carbon series, C18:0, C18:1, C18:2, and C18:3 represent stearic, oleic, linoleic, and linolenic acids, respectively. One or two letter abbreviations also are used, and these four acids sometimes are designated by St, O, L, and Ln, respectively.

Under the most common chemical convention, fatty acids are named on the basis of the number of carbon atoms, starting with the terminal carboxyl (COOH) carbon as number 1. The terminal letter *e* of the respective alkane hydrocarbon is replaced with *oic* to indicate an acid:

$$CH_3CH_2CH_2CH_2CH_2CH_2CH_2CH_3$$

Octane

$$\overset{8}{CH_3}\overset{7}{CH_2}\overset{6}{CH_2}\overset{5}{CH_2}\overset{4}{CH_2}\overset{3}{CH_2}\overset{2}{CH_2}\overset{1}{COOH}$$

Octanoic acid

The suffix *dioic* is used if the acid contains two carboxyl groups.

When double bonds are present, the suffix *anoic* is changed to *enoic*, *dienoic*, or *trienoic* to indicate the number of bonds present, and the location of the first carbon in the double bond is indicated by a number preceding the systemic name. Under this convention, stearic, oleic, linoleic, and linolenic acids are called octadecanoic, 9-octadecenoic, 9,12-octadecadienoic, and 9,12,15-octadecatrienoic acids, respectively.

The geometric configuration of the double bonds is indicated by the latin prefixes *cis*- (both hydrogens on one side) and *trans*- (hydrogens across from each other). Linoleic acid, with *cis* configuration in both double bonds, is called *cis*-9,*cis*-12-octadecadienoic acid. Most fatty acids occur in nature in the *cis* form. Oleic acid is automatically assumed to be in the *cis* configuration, and the corresponding *trans* form is called elaidic acid.

cis *trans*

The same system (with the carboxyl carbon designated as 1) also is used in naming fatty acids that are branched or include unusual structures, with numbers indicating the carbon to which the respective group or structure is attached.

Occasionally in the literature, the carboxyl unit is regarded as a group substituted for hydrogen. In this case, the number 1 location is moved one position away from the reactive end, and the suffix *carboxylic acid* is added:

$$\overset{5}{C}H_3\overset{4}{C}H_2\overset{3}{C}H_2\overset{2}{C}H_2\overset{1}{C}H_2COOH$$

1-Pentanecarboxylic acid
(hexanoic acid)
(caproic acid)

Locations of the last double bond are of special interest in long-chain polyunsaturated fatty acids. Their numbering sometimes begins from the noncarboxyl end, with the methyl carbon counted as 1 and the first unsaturated carbon designated as an *n* or omega (ω) number. Thus, linoleic acid (9,12-octadecadienoic) also may be designated as C18:2ω6. The polyunsaturated fatty acids consist mainly of two groups whose first double bonds occur either three or six carbons from the methyl end (n-3, ω-3 or n-6, ω-6, respectively). Distributions of individual fatty acids can vary by over 100 percent in the oil of a species. Table 8.2 summarizes general distribution for the major edible fats and oils, and Table 8.3 that for industrial fats and oils.

In recent years, research attention has turned to evaluating the potential of nontraditional fats/oils sources, including newly domesticated crops, forest oilseeds, bacteria, yeasts and molds, and algae. The reader should consult the references for types of fatty acids and their occurrence in these sources.[3,4]

Acylglycerols

Glycerol esterified with one, two, or three fatty acids is found in nature and can be made commercially. The designations monoacylglycerol, diacylglycerol, and triacylglycerol, respectively, now are encouraged in the scientific literature although the older terms mono-, di-, and triglycerides still are widely used in commerce. As the number of fatty esters on the glycerine decreases, the compound becomes more polar and functionally effective as a surfactant. Mono- and diacylglycerols are described in more detail in the section on emulsifiers.

The triacylglycerols are named in various ways. For example, unsaturated fatty acids sometimes are indicated as *U* and the saturated as *S*. If glycerol is completely esterified with stearic acid, the resulting monoacid triacylglycerol may be designated as SSS, or, more descriptively as StStSt, tristearin, tristearoylglycerol, or glycerol tristearate.

If more than one species of fatty acid is present, its relative location on the glycerol may be important to its functionality, enzyme susceptibility, and storage stability of the fat/oil. Several conventions have been developed to specify arrangements of fatty acids on the glycerol molecule (if known). To avoid confusion from inversion of the 1 and 3 carbon positions, hierarchies have been established to designate the number 1 carbon under the R/S system[5] and the *sn* (stereospecific numbering) system.[6] Three abbreviations are used under the *sn* system to indicate location of specific fatty acids in triacylglycerol molecules: *sn* immediately before the word glycerol, as in 1-stearoyl-2-oleoyl-3-myristoyl-*sn*-glycerol, indicates that the respective fatty

TABLE 8.1 Alternative Names and Selected Characteristics of Some Important Fatty Acids

Carbon Atoms	*Common Name*	*Letter Abbr.*	*Systematic Name*	*Melting Point, °C*	*Iodine Value*	*Common Sources*
Saturated Fatty Acids						
3	Propionic	—	Propanoic	−20.8	—	Bacterial fermentation
4	Butyric	—	Butanoic	−7.9	—	Milk fats
5	Valeric	—	Pentanoic	−33.8	—	Bacterial fermentation
5	Isovaleric	—	3-Methylbutanoic	−51.0	—	Dolphin and porpoise fats
6	Caproic	—	Hexanoic	−3.4	—	Milk fats, some seed oils
8	Caprylic	—	Octanoic	16.7	—	Milk fats, *Palmae* seed oils
10	Capric	—	Decanoic	31.6	—	Ship and goat milk, palm seed oils, sperm head oil
12	Lauric	—	Dodecanoic	44.2	—	Coconut-oil
14	Myristic	—	Tetradecanoic	54.4	—	Palm and coconut oils
16	Palmitic	—	Hexadecanoic	62.9	—	Palm oil
18	Stearic	St	Octadecanoic	69.6	—	Animal fats
19	Tuberculostearic	—	10-Methylstearic	11.0	—	Tubercle bacillus lipids
20	Arachidic	—	Eicosanoic	75.4	—	Some animal fats
22	Behenic	—	Docosanoic	80.0	—	Peanut and various other seed oils
24	Lignoceric	—	Tetracosanoic	84.2	—	Minor amounts in some seed oils
26	Cerotic	—	Hexacosanoic	87.8	—	Plant waxes
28	Montanic	—	Octacosanoic	90.9	—	Beeswax and other waxes
30	Mellisic	—	Triacontanoic	93.6	—	Beeswax and other waxes
Unsaturated Fatty Acids						
10	Caproleic	—	9-Decenoic	—	149.1	Milk fats
10	Stillingic	—	2,4-Decadienoic	—	—	*Stillingia* oil
12	Lauroleic	—	2-Dodecenoic	—	128.0	Butterfat

14	Myristoleic	—	9-Tetradecenoic	18.5	112.1	Some feed fats, milk fats
16	Palmitoleic	—	9-Hexadecenoic	—	99.8	Many fats and marine oils
16	Hiragonic	—	6,10,14-Hexadecatrienoic	—	—	Sardine oil
17	—	—	9-Heptadecenoic	14.0	—	*Candida tropicallis* yeast
18	Elaidic	—	9-Octadecenoic	43.7	—	Butterfat
18	Oleic	—	9-Octadecenoic	16.3	89.9	Almost all fats and oils
18	Petroselinic	—	6-Octadecenoic	30–33	—	Parsley seed oil
18	Vaccenic	—	11-Octadecenoic	44.0	—	Butterfat
18	Linoleic	Lo	9,12-Octadecadienoic	−6.5	181.0	Most vegetable oils
18	—	—	6,9,12-Octadecatrienoic	—	—	Evening primrose family
18	Eleostearic	—	9,11,13-Octadecatrienoic	—	—	Tung oil
18	Linolenic	Ln	9,12,15-Octadecatrienoic	−12.8	273.5	Linseed oil, soybean oil and canola oil
20	Gadoleic	—	9-Eicosenoic	23–24	81.8	Some fish oils
20	—	—	5,8,11-Eicosatrienoic	—	—	Brain phospholipids
20	—	—	8,11,14-Eicosatrienoic	—	—	Shark liver oil
22	Erucic	—	12-Docosenoic	33.5	75.0	Rapeseed oil
22	—	—	13,16-Docosadienoic	—	—	Rapeseed oil
Fatty Acids of More Unusual Structure						
18	Chaulmoogric	—	13,(2-Cyclopentenyl)tridecanoic	68.5	90.5	*Chaulmoogra* oil
18	Malvalic	—	8,9-Methylene-8-heptadecenoic	—	—	*Malvaceae* seeds, cottonseed
18	Ricinoleic	—	12-Hydroxy-9-octadecenoic	5.5	86.0	Castor oil
18	Vernolic	—	12,13-Epoxy-9-octadecenoic	30–31	—	Some *Compositae* seeds
19	Sterculic	—	9,10-Methylene-9-octadecenoic	—	—	*Sterculiaceae* seeds, cottonseed
20	Arachidonic	—	5,8,11,14-Eicosatetraenoic	−49.5	333.5	Lard
20	—	—	5,8,11,14,17-Eicosapentaenoic	—	—	Some fish oil
20	Lesquerolic	—	14-Hydroxy-11-eicosenoic	—	—	*Lesquerella* seed oil
22	—	—	4,7,10,13,16,19-Docosahexaenoic	—	—	Some fish oil

TABLE 8.2 Fatty Acid Composition of Some Edible Oils and Fats[a]

Source	*<14:0*	*14:0*	*16:0*	*16:1*	*18:0*	*18:1*	*18:2*	*18:3*	*20:0*	*20:1*	*22:0*	*22:1*	*24:0*	*24:1*
Almond oil	—	0.0	6.5	0.6	1.7	69.4	17.4	—	—	—	—	—	—	—
Avocado oil	—	—	11.0	3.4	0.7	71.5	12.0	1.5	—	—	—	—	—	—
Barley bran oil	—	0.5	10.8	0.2	1.0	17.8	55.3	4.4	—	—	—	—	—	—
Borage oil	—	—	11.3	—	3.7	16.3	38.1	23.0	0.2	3.9	—	2.4	—	1.4
Buffalo gourd seed oil	—	—	11.8	—	3.5	21.9	60.6	0.0	0.0	—	—	—	—	—
Butter fat	23.8	8.2	21.3	1.8	9.8	20.4	1.8	1.2	—	—	—	—	—	—
Canola oil[b]	—	—	4.8	0.5	1.6	53.8	22.1	11.1	1.1	1.5	0.3	0.1	0.1	—
Cherry pit oil	—	—	7.8	0.4	2.4	43.9	44.8	0.5	0.7	—	—	—	—	—
Cocoa butter	—	0.1	25.4	0.2	33.2	32.6	2.8	0.1	—	0.0	—	—	—	—
Coconut oil	58.7	16.8	8.2	—	2.8	5.8	1.8	—	—	—	—	—	—	—
Corn oil	0.0	0.0	10.9	—	1.8	24.2	58.0	0.7	—	—	—	—	—	—
Cottonseed oil	—	0.8	22.7	0.8	2.3	17.0	51.5	0.2	—	—	—	—	—	—
Evening primrose oil	—	—	8.5	—	2.5	8.5	72.5	11.0	—	—	—	—	—	—
Fish (manhaden) oil	—	9.6	20.5	12.6	3.3	11.0	0.7	1.6	0.3	—	—	0.8	—	—
Grapeseed oil	—	0.1	6.7	0.3	2.7	15.8	69.6	0.1	—	—	—	—	—	—
Illipe butter	—	—	23.7	—	19.3	43.3	13.7	—	—	—	—	—	—	—
Lard	0.5	1.3	23.8	2.7	13.5	41.2	10.2	1.0	—	1.0	—	—	—	—
Lupine oil	—	—	8.3	—	2.5	55.0	17.7	9.3	—	—	—	—	—	—
Macademia nut oil	—	0.6	8.5	21.7	3.7	56.0	1.7	—	—	1.4	—	—	—	—
Mango kernel oil	—	—	7.6	—	36.0	49.4	5.0	0.5	1.4	—	—	—	—	—
Mustard seed oil	—	0.1	1.9	0.3	0.1	17.7	9.1	0.5	0.6	3.91	1.8	55.1	0.2	1.9
Okra seed oil	—	0.2	33.7	0.6	3.3	17.9	42.2	0.2	0.1	—	0.2	—	—	—
Olive oil	—	0.0	11.0	0.8	2.2	72.5	7.9	0.6	—	—	—	—	—	—
Palm oil	0.1	1.0	43.5	0.3	4.3	36.6	9.1	0.2	—	0.1	—	—	—	—
Palm kernel oil	54.2	16.4	8.1	—	2.8	11.4	1.6	—	—	—	—	—	—	—
Peanut oil	—	0.1	9.5	0.1	2.2	44.8	32.0	—	—	1.3	—	—	1.8	—
Rapeseed oil[c]	—	—	1.7	—	0.9	12.3	12.7	7.6	1.2	5.8	0.9	59.4	0.5	1.6
Rice bran oil	—	0.7	16.9	0.2	1.6	39.1	33.4	1.6	—	—	—	—	—	—
Safflower oil	—	0.1	6.2	0.4	2.2	11.7	74.1	0.4	—	—	—	—	—	—
Safflower oil[d]	—	—	4.8	—	1.3	75.3	14.2	—	—	—	—	—	—	—
Sal seed oil	—	—	5.3	—	34.0	49.1	3.8	3.3	4.0	—	—	—	—	—
Sesame oil	—	—	8.9	0.2	4.8	39.3	41.3	0.3	—	0.2	—	—	—	—
Shea butter	1.7	0.1	4.4	0.1	38.8	43.5	4.9	0.3	—	0.0	—	—	—	—
Soybean oil	—	0.1	10.3	0.2	3.8	22.3	51.0	6.8	—	—	—	—	—	—
Sunflower oil[e]	—	—	5.4	0.2	3.5	80.6	8.4	0.2	0.3	—	—	—	—	—
Sunflower oil[f]	—	0.1	5.9	0.1	4.5	19.5	65.7	—	—	—	—	—	—	—
Tallow	0.9	3.7	24.9	4.2	18.9	36.0	3.1	0.6	—	0.3	—	—	—	—
Teaseed oil	0.1	0.1	17.5	0.5	3.1	49.9	22.2	0.7	—	1.0	—	—	—	—
Tomato seed oil	—	0.2	15.0	0.5	4.4	21.9	50.8	2.3	—	—	—	—	—	—
Walnut oil	—	—	7.0	0.1	2.0	22.2	0.4	52.9	10.4	—	—	—	—	—
Wild cucurbit oil	—	—	19.0	—	—	34.0	47.0	15.0	—	—	—	—	—	—

[a]These are average values from recent years' crops. [b]Low-erucic-acid variety rapeseed. [c]High-erucic-acid variety. [d]High-oleic variety. [e]Grown in colder climates. [f]High-oleic variety.

acids are in the 1, 2, and 3 positions; the term *rac* (racemic mixture), as in rac-StOM, indicates that the middle acid is in the glycerol 2 position, and the remaining fatty acids are equally divided between the sn-1 and sn-3 positions. The term β, as in β-StOM, indicates that the middle acid is in the glycerol 2 position, but distribution of the other two acids is unknown.[7]

Melting Points

Factors affecting melting points of different fat samples include the types of fatty acids

TABLE 8.3 Fatty Acid Composition of Some Industrial Oils and Fats[a]

Source	*<14:0*	*14:0*	*16:0*	*16:1*	*18:0*	*18:1*	*18:2*	*18:3*	*20:0*	*20:1*	*22:0*	*22:1*	*22:2*	*24:1*
Castor oil[b]	—	—	1.1	0.2	1.0	3.3	3.6	0.32	0.4	—	—	—	—	—
Chinese tallow	1.3	2.1	65.0	—	4.4	22.5	0.8	—	—	—	—	—	—	—
Crambe oil	—	—	2.0	0.4	0.4	16.9	8.6	6.4	0.5	3.2	2.0	57.2	—	1.4
Crepsis foetida oil[c]	—	0.1	4.8	—	2.9	4.3	27.8	—	—	0.3	—	—	—	—
Croton oil	2.5	5.4	6.2	0.2	3.2	15.8	49.4	3.0	2.9	8.9	0.2	0.6	—	—
Cuphea oil	—	76.4	7.8	2.4	—	0.7	5.9	6.9	0.1	—	—	—	—	—
Jojoba oil	—	—	1.0	—	—	9.0	—	—	—	70.7	—	16.3	—	3.0
Lesquerella seed oil	—	1.4	1.4	2.0	17.6	7.9	13.2	—	52.5	—	—	—	2.7	—
Linseed oil	—	—	5.3	—	4.1	20.2	12.7	53.3	—	—	—	—	—	—
Meadowfoam oil	—	—	—	—	—	—	—	—	—	64.5	—	18.5	13.5	—
Neatsfoot oil	—	0.7	16.9	—	2.7	64.4	2.3	0.7	0.1	—	—	—	—	—
Oitica oil[d]	—	—	7.0	—	5.0	6.0	—	—	—	—	—	—	—	—
Rapeseed[e]	—	0.1	2.6	0.3	0.9	11.2	12.8	8.6	—	7.5	—	48.1	—	—
Rubber seed oil	—	0.2	19.1	—	17.8	24.5	30.5	2.4	—	0.1	0.9	0.4	—	—
Stokes aster oil[f]	—	—	2.8	—	0.9	7.0	16.5	—	—	—	—	—	—	—
Tall oil	—	—	—	—	—	50.0	7.0	41.0	—	—	—	—	—	—
Tung oil	—	—	3.1	—	2.1	11.2	14.6	69.0	—	—	—	—	—	—
Veronia seed oil[g]	—	—	2.7	—	1.3	2.0	8.8	0.4	—	—	—	—	—	—
Whale oil	—	3.3	8.1	26.9	1.1	33.3	—	—	—	10.9	—	2.2	—	—

[a]These are average values from recent years' crops. [b]Contains 89.2% ricinoleic and 1.4% dihydroxystearic acids. [c]Contains 59.8% crepenynic acid. [d]Contains 78.0% licanic acid and 4.0% hydroxy acids. [e]High-erucic-acid variety. [f]Contains 71.3% venolic acid. [g]Contains 78.5% vernolic acid and 5.8% hydroxy fatty acids.

present in the triacylglycerol and their location, chain length of the fatty acids, number and location of *cis* and *trans* double bonds on the fatty acid chains, and compatibility of the different triacylglycerols in the mixture and the type of crystal present.

Among the fatty acids, melting points increase with chain length. *Trans* fatty acids always have higher melting points than their *cis* counterparts for any chain length. Where only one double bond exists in a fatty acid, as in C18:1, the melting point is lower if it is located after an odd-number carbon than an even-number carbon, and also if the double bond is located near the middle of the chain as compared to a location at either end.[8] Compatibility of mixed fats from different sources also can be a factor. For example, the melting point of mixed triacylglycerols consisting primarily of 16 to 18 carbon fatty acids, whether of animal or vegetable origin, generally rises smoothly with increased content of higher melting fats. However, when triacylglycerols with 12 to 16 carbon fatty acid fats (from coconut, palm kernel, or palm oil) are added, concentration–related eutectic points and regions of noncompatibility (manifested as mixtures of solids in oil) may occur.[9,10]

Polymorphism and Crystal Types

The same sample of a pure or mixed fatty acid triacylglycerol may show as many as five different melting points, depending on its heating and cooling history. For example, tristearin has an amorphic (noncrystalline) solid form and three crystalline forms with melting points at 54.7°C, 63.2°C, and 73.5°C. If a small sample of fat is heated slowly (as in a capillary tube melting point determination), melting, resolidification, and remelting at increasingly higher temperatures may be observed. These changes result from the fat seeking the most compact crystal form and the most stable (lowest) thermodynamic energy state possible.

As fatty acids or triacylglycerols cool, Gibbs free energy ($G = H - TS$) decreases by reduction of both enthalpy (H) and entropy (S, the degree of disorder). At decreased S, the fatty acid chains assume polelike structures

that are less co-repulsive and pack more tightly into crystal lattices. It is generally accepted that, to participate in a crystal structure, triacylglycerols assume an "h" configuration (also called a two-legged chair or tuning fork). This can be envisioned in our gravity-oriented macro world by assuming that each ester linkage acts as a hinge. If the glycerol number 2 carbon chain is held upright at the methyl end, the number 3 carbon chain falls directly beneath it, forming the back and one leg of the chair, and the number 1 carbon chain juts out at a right angle and curves downward to form the second leg.[11-13] In order to save space, half of the chairs are packed upside down to form a palisade-like structure. A stack of two such structures forms a bilayer whose outer surfaces consist of methyl groups, displaying low interattraction. This helps explain why triacylglycerol crystals are relatively flat, grow rapidly in length and less rapidly in width by adding parallel chair structures, but grow slowly in thickness by adding additional bilayers.

Within each layer, the chair backs and legs can be further envisioned to act like a vertical loose palisade of bumpy-surfaced posts. (Although the hydrocarbon chains have stiffened into pole shape, the carbon atoms are not positioned as beads on a tight string, but rather in sawtooth-like fashion with carbon–carbon bonds of 112°). The posts can slide behind each other to obtain a tighter packing with lowered free energy, and can be tilted in two directions to allow the sawtooth-configured carbon atoms on adjacent chains to slip by each other and pack more tightly. Also, the vertical units in one layer can all be tilted at an opposing angle to the units in the other layer.

On cooling, fats may assume various polymorphic forms. On slow cooling, or reheating of the vitreous form, they change to crystalline forms. The form showing the least amount of crystalline order for a triacylglycerol, as determined by X-ray diffraction and infrared spectroscopy, is called the *alpha* (α) form. The most compact crystalline form, with the lowest free energy and the highest melting point, is called the *beta* (β) form. One or more intermediate *beta prime* (β') forms also may exist, and are indicated as β_3', β_2', and β_1' as the crystal progressively assumes tighter packing with lower free energy and increased melting points.[14]

When a monoacid triacylglycerol is cooled slowly and without mixing, it preferentially assumes the β crystal form. However, the required tight packing of crystals requires time for alignment, and may be thwarted by increased viscosity as the fat/oil mixture cools. However, if the heated oil is cooled rapidly, it may first be rushed into an amorphous (vitreous, glasslike) sub-α state. Upon exposure to heat, the fat then may pass directly to the β crystal form, or stepwise through the α and β' forms until it reaches the β form.

The free energy relationships between the different crystal forms are depicted in Fig. 8.1. Although the β crystal form has the lowest free energy G, its formation also has the highest activation energy ΔG. Differential scanning calorimetry (DSC) often is used to follow free energy changes as fats melt and change between their polymorphic forms.

Formation of β crystals may intentionally be encouraged or kinetically hindered depending on the specific application. The α crystals are relatively unstable, and commercial interest is primarily placed on the differences between the β' and β forms. Generally, β' crystals are

Fig. 8.1. Gibbs free energy relationships for establishment of α, β', and β crystals. (*By permission of Marcel Dekker, Inc., New York.*)

Fig. 8.2. Drawings of two- and three-chain triacylglycerol layers: (a) saturated monoacid SSS-type; (b) saturated symmetrical PSP-type where 2 chain differs from 1 and 3 chains in length; (c) symmetrical POP-type where 2 chain is unsaturated; and (d) symmetrical OPO-type where 1 and 3 chains are unsaturated. Crystals grow in bilayer units.

smaller (about 0.5–2.0 diameter microns in shortening, and 5–10 microns in margarine),[15] whereas β crystals can grow to as large as 20 to 30 microns. When the objective is to thermally fractionate fats by crystallization, production of the β form is encouraged by carefully controlling temperatures (to not shock the fat into a lower melting β' form), gentle stirring, and nuclei seeding. The smaller β' crystals have smoother mouthfeel, minimize oiling off of margarine, and entrap more air in creaming cake batters. Their production is intentionally encouraged by formulating mixtures of natural or preprocessed fats, by inclusion of emulsifiers to interfere with crystal growth, and by rapid agitation during plasticizing of the margarine or shortening.

An awareness of crystal packing characteristics and polymorphism helps one to understand incompatibility problems of different fats. As shown in Fig. 8.2, if the three fatty acids in the triacylglycerol are saturated and approximately of the same length, each of the layers in the fat bilayer will be approximately two fatty acid chains in height. If the fatty acid on the number 2 glycerol carbon is appreciably shorter than those on the 1 and 3 carbons, each of the layers will be three chains thick. If the triacylglycerol is symmetrical, with either the number 2 glycerol carbon or both the 1 and 3 carbons unsaturated, the layer will be three chains long, but with a zig-zag configuration to accommodate the *cis* configuration.

Crystal formation has specific demands, and individual crystals in mixed systems consist of only one species of triacylglycerol. However, surfactants and other molecules can act as impurities and interrupt crystal growth. Different triacylglycerols are considered *compatable* when they co-crystalize as separate crystals under the same conditions without formation of a eutectic.

Waxes

Waxes are fatty acid esters of alchols and are formed by the general reaction:

$$\underset{\text{Alcohol}}{CH_3(CH_2)_nCH_2OH} + \underset{\text{Fatty acid}}{CH_3(CH_2)_nCOOH} =$$

$$\underset{\text{Wax ester}}{CH_3(CH_2)_nCH_2COCH_2(CH_2)_nCH_3} + H_2O$$

A major role of waxes in nature is protection of plant tissues. Examples include coating upper surfaces of leaves to reduce dehydration by the sun and protecting seeds against moisture loss during storage. Minor quantities of waxes always are extracted from oilseeds by the commercial hexane process. Waxes extracted from seed hulls have been a problem in sunflower oil because their presence shortens

TABLE 8.4 Sources and Compositions of Natural Waxes

Type	*Melting Point °C*	*Main components*
Animal waxes		
Beeswax	64	Myricyl palmitate
Chinese	82–84	Isoheptacosyl isoheptacosanoate, ceryl lignocerate
Shellac	81–82	Ceryl lignocerate, ceryl cerotate
Spermaceti	—	Cetyl palmitate
Wool (anhydrous lanolin)	36–42	Cholesteryl estolidic esters, alcohol esters of iso- and anteiso acids
Mineral waxes		
Montan	86	Tricontanyl esters of C_{28-30} acids
Petroleum waxes		
Microcrystalline	71–88	Hydrocarbons (490–800 molecular weights)
Paraffin	54–57	Hydrocarbons (350–420 molecular weights)
Vegetable waxes		
Bayberry	43–48	Trimyristin, tristearin
Candelilla	70–80	C_{29-33} hydrocarbons, simple esters and lactones
Carnauba	80–85	Esters of C_{26-30} alcohols and C_{26-30} ω-hydroxy acids
Esparto	69–81	Hydrocarbons, esters of C_{26-32} acids and alcohols
Japan	51–62	Tripalmitin
Jojoba (a liquid wax)	11–12	Docosenyl eicosanoate
Ouricury	79–85	Myricyl cerotate and hydroxycerotate
Sugarcane	79–81	Myricyl palmitate stigmasteryl palmitate

the cold test of refrigerated oil. Sunflower oil may be dewaxed by first degumming or miscella-refining to remove the natural emulsifier lecithin, which limits growth of wax crystals, and then winterizing (chilling and filtering) the oil or its miscella.[16,17] With improvement of dehulling equipment, some sunflower seed processors have begun to remove the hulls before extraction. Domestic interest only now is turning to commercial extraction (*oil milling*) of rice bran, whose oil is believed to contain waxes that have commercial promise.

In some species, waxes serve the function of high-energy storage forms. Waxes are common in the oils of fish and other marine animals. The major lipids of commercial whale oil consist of approximately 65 percent waxes and 35 percent triacylglycerols. Essentially all of the oil in jojoba (*Simmondsia chiensis*) seed is in wax form. Whale and jojoba oils have been valued for stability in heavy-duty lubrication applications, and jojoba oil has been used in a variety of cosmetics.

Extraction processes for waxes vary in sophistication, from boiling and skimming crushed leaves and berries—as in the production of Candelilla wax (a substitute for hard carnauba wax) from *Euphorbia antisyphilitica* in the Big Bend areas of Texas and Mexico—to screwpressing seeds such as those of jojoba. Cold-pressed jojoba oil is preferred for cosmetics because of concern about other lipids that might be extracted in solvent processes, and prepressed jojoba meal sometimes is extracted by a secondary hexane process to recover oil for industrial uses. Waxes are susceptible to hydrolysis by nonspecific lipases, and also are at least partially digestible. They may be hydrogenated or sulfurized into solid forms. Compositions of significant commercial waxes from natural sources are given in Table 8.4

Terpenes, Carotenoids, Fat-Soluble Vitamins, and Steroids

Terpenes are condensation products of the five-carbon hydrocarbon *isoprene* (2-methyl-1,3-butadiene), and are extractable by nonpolar solvents. They are classified according to the number of isoprene units: two units, *mono-*

terpenes; three units, *sesquiterpenes*; four units, *diterpenes*; six units, *triterpenes*; eight units, *tetraterpenes*; and polyterpenes. Terpenes may be linear or cyclic. Taken together, this class of compounds includes the major essential oils, fat-soluble colors, fat-soluble vitamins, and steroids. A short-hand (sawtooth) form often is used to depict the longer chains.

Examples of monoterenes include the linear aldehyde *citral* which is found in many essential oils, and the (*cis*) alcohol *geraniol*—a major component of oil of geranium. Cyclic mono-terpenes include *limonene*, *menthol*, *pinene*, *camphor*, and *carvone*—major components of lemon oil, mint oil, turpentine, camphor oil, and caraway oil, respectively. Sesquiterpenes include *farnesol*, a component of rose oil, and *bisabolene*, a component of Bisabol myrrh. The diterpenes include *phytol*, a component of chlorophyll, and vitamin A, which is one-half of the tetraterpene *β*-carotene. The triterenes include *squalene*, a precursor of cholesterol. Examples of tetraterpenes are the oil-soluble *carotenoid* plant pigments: xanthophylls, including the yellow pigments lutein in plant leaves and zeaxanthin in corn (*Zea mays*); capsanthin, the red pigment in red peppers (*Capsicum annum*); lycopene, the red coloring of tomatoes (*Lycopersicum esculentum*); and *β-carotene*, a yellow-orange pigment that is the precursor of Vitamin A. *Bixin*, from the seed pods of *Bixa orellana*, is the yellow-orange pigment in annatto food color; it is considered to be a form of carotene oxidized to remove both six-membered end rings. Polyterpenes include gutta, a natural water-repellant and electrical insulating material from *Palaquium gutta*, and natural rubbers with molecular weights of up to 1.2 million obtained from the latex of the *Hevea brasiliensis* tree and from rubber-filled cells in the branches and roots of the guayule shrub (*Parthenium argentatum*).

Important steroids derived from squalene include cholesterol, stigmasterol, sitosterol, and campesterol.

Isoprene

Bixin

Geraniol

β-Carotene

Limonene

Zeaxanthin

Squalene

Lycopene

Cholesterol

Stigmasterol

α_1-Sitosterol

β-Sitosterol

Campesterol

The fat-soluble vitamins include: vitamin A (*retinol*), a colorless compound considered to result from cleavage of β-carotene; vitamin D_3 (*cholecalciferol*), a steroid; and vitamins E and K (*α-tocopherol* and *menadione and its derivatives*, respectively), which consist of isoprene and phenolic-quinone components.

Vitamin A

Vitamin D_3

Vitamin E

Vitamin K_1

Vitamin K_3

The carotenoids generally are highly susceptible to oxidation. Only selected members (xanthophylls, β-carotene, and bixin) are extracted for food coloring, and then they are used in products such as butter, margarine, and cheese, where exposure to oxygen and light is limited. The more stable pigments (lutein and xanthophylls) are extracted from alfalfa and marigold blossoms for use as poultry feed additives to increase yellow color in egg yolks and in the skins of broilers. Much of the world's β-carotene is produced synthetically, but crude (red) palm oil may be a significant future source.

Oils from livers of cod and other fish species were early sources of vitamin D_3; also, ergosterol was irradiated by ultraviolet light to produce vitamin D_2. Currently, most of the commercial vitamin D also is produced synthetically. Vitamin E has been obtained commercially from extracted wheat germ oil, and recovered from deodorizer condensates from oil refining plants. Palm oil refining residues also are indicated major sources in the future.

SOURCES AND UTILIZATION OF FATS AND OILS

Production

Over 50 percent of the world's production of vegetable oils enters international trade, 19 percent in seed form and 33 percent as extracted oil. Extensive records are kept, country by country, of production, imports, and exports. Global production and export data for 1987/88 are shown in Table 8.5. These figures include crops harvested in the last half of the first year in the Northern Hemisphere and during the first half of the succeeding year in the Southern Hemisphere.

By far, soybean is the world's largest source of visible fats/oils, accounting for approximately 21.5 percent of total production. It is followed by palm oil, rapeseed/canola oil, sunflower seed oil, and tallow and grease in tonnages produced. The United States is the world's leader in soybean growing, accounting for approximately 44 percent of the total tonnage produced. It is followed by Brazil, Argentina,

TABLE 8.5 Area, Yield, Oilseeds, Products, and Exports of World's Major Edible-Type Oils for 1987/88[1]

Oil/Fat Source	*Area*[a]	*Yield*[b]	*Oilseed*[c]	*Meal*[d]	*Oil*[e]	*Exports*[f]
Edible Fats						
Soybeans	53.48	1.82	103.66	67.55	15.24	3.84
Cottonseed	32.19	0.97	31.31	9.12	3.48	0.39
Peanut	18.38	1.13	20.86	4.98	3.10	0.31
Sunflower seed	14.95	1.31	20.92	7.75	7.28	2.18
Rapeseed/canola	15.55	1.30	23.44	9.05	7.75	1.83
Flaxseed/linseed	4.18	0.54	2.27	0.86	—	—
Copra/coconut	—	—	4.35	1.55	2.68	1.38
Palm oil	—	—	—	—	8.36	5.55
Palm kernel	—	—	2.69	1.43	1.20	0.74
Olive	—	—	—	—	1.90	0.46
Fish	—	—	—	6.79	—	—
Butter[g]	—	—	—	—	5.30	—
Total Edible	138.73	7.07	209.51	109.08	56.29	16.78
Industrial Fats						
Industrial linseed	—	—	—	—	0.62	0.22
Marine oils fish	—	—	—	—	1.54	0.76
Tallow and grease[h]	—	—	—	—	6.56	—
Total Industrial	—	—	—	—	8.72	0.98
Total Fats and Oils	138.73	7.07	209.51	109.08	65.01	17.76

[1]USDA FAS Circular Series FOP 11-90, Nov. 1990.
[a]In million hectares (2.47 acres). [b]In metric tons per hectare. [c]In million metric tons. [d]In million metric tons, converted to 44% (soybean) protein equivalent. [e]In million metric tons. [f]In million metric tons, transhipments for palm oil through Singapore excluded. [g]Converted to fat content basis. [h]Inedible and edible tallows and greases included.

and China. In 1987/88, 41 percent of United States produced soybean oil was exported, compared to 15 percent and 25 percent, respectively, for Brazil and approximately 22 percent and 90 percent, respectively, for Argentina. These two countries impose larger severance taxes on exports of raw crops to encourage exports of value-added products, thus keeping the processing jobs at home. Several countries, notably The Netherlands, Portugal, Taiwan, and Singapore, have established significant process-in-transit industries, importing whole soybeans and exporting the meal and oil.

The leading producers of other oilseeds are: for *rapeseed/canola*—China, the EC-12 (European Common Market), India, Canada, and Eastern European countries; for *sunflower seed*—Russia, Argentina, the EC-12, and Eastern European countries: for *ground nuts, peanuts*—India, China, and the United States; for *cottonseed*—China, Russia, India, the United States, and Pakistan; and for *palm oil*—Malaysia and Indonesia, with oil palm plantations and oil mills coming into production in many equatorial countries.

Many factors dictate which oilseed species will be grown and/or imported into a country. The local need for high-protein feed meals probably is the leading factor. As shown in Table 8.6, the oil content of row crop oilseeds varies from about 20 percent for soybeans to 45 percent for sunflower seed and 43 to 45 percent for rapeseed/canola. More feed coproducts always are produced than oils—by a ratio of as much as 4:1 in the case of soybeans. Soybean meal is the major feed protein source for production of poultry, currently the leading domestic and global meat source, and for the rapidly developing aquaculture industries. For example, Malaysia, which produces nearly two-thirds of the

TABLE 8.6 Gross Composition of Major Edible Oilseeds

Crop/Source	*Moisture*	*Protein*	*Fat*	*Fiber*	*Ash*	*Oil:Co-products Ratio*
Soybean	4.2	37.0	24.0	3.6	4.3	1:3
Cottonseed	4.7	32.6	36.3	2.0	4.6	1:2
Peanut	6.7	25.7	49.2	4.9	2.3	1:1
Sunflower	5.4	22.8	49.6	4.2	3.5	1:1
Safflower	5.6	16.2	38.5	2.5	5.5	1:2
Copra	3.0	6.9	64.5	5.3	1.9	2:1
Canola	7.5	21.8	44.9	4.5	4.7	1:1
Palm kernel	6.5	8.3	49.0	8.1	2.0	1:1
Sesame	4.7	17.7	49.7	4.6	4.5	1:1

world's exported palm oil, imports soybeans for crushing, retains the meal for local animal feeds industries, and exports the oil.

Cool weather and short growing seasons have essentially limited Canada, Northern Europe, and the European Russian republics to growing rapeseed/canola or sunflower seed. Insect problems were important factors in selecting crops before modern insecticides and integrated pest management systems became available. Devastation of the southeastern United States cotton crop by the boll weevil led to the introduction of peanut growing and to initial domestic acreages of soybeans as alternative crops. Also, subsidies to local farmers, by governments that have chosen self-sufficiency in food production in preference to free trade, have encouraged the growing of crops that otherwise would not be profitable in some climates.

The relative availability of fats/oils from by-products of other current agribusiness and commercial enterprises is an additional consideration in the types and amounts of oilseeds grown. As examples, cotton lint accounts for approximately 85 percent of total returns to the cotton farmer, with only 15 percent coming from the seed. The domestic supply of cottonseed for crushing is dictated by world demand and price of cotton and by whole-seed feed needs of the southern United States dairy industry. Availability of cottonseed for crushing is relatively unresponsive to the price of cottonseed oil. Although soybeans are the major supply of edible oil, their production is closely related to the demand for soybean meal, which typically accounts for 67 to 75 percent of the market value of this crop. Corn oil has become the second major oil in the United States last decade owing to the large quantities of corn germ made available by rapid growth of the domestic corn sweetener industry. Beef, pork, and poultry packing operations always produce fatty tissues for rendering into edible tallow, lard, and chicken fat, respectively. Generally, fats/oils from coproducts of other local processing industries must clear the market first, at whatever the going price, before growing or importing of significant quantities of high-oil-content crops such as rapeseed/canola, sunflower seed, and peanuts becomes economically attractive.

Consumption

Average per capita consumption of fats/oils is difficult to determine from gross disappearance figures because these materials also are used in animal feeds and industrial applications, but it generally is related to personal income. Consumption of oils is estimated at 10.4 kg/person for the world, and ranges from 27.2, 42.9, and 29.3 kg for the United States, Belgium, and West Germany, respectively, to 13.1, 4.7, and 5.9 kg, respectively, for Egypt, China, and India. Approximately 60 percent of the total fat consumed domestically is invisible in meat, poultry, fish, dairy products, and eggs, and the 40 percent visible fat is used primarily in the form of salad and cooking oils, shortenings, and margarine. The world needs about 1.5 to 2.0 million additional

tons of fats/oils each year for its rapidly growing population.

From 1965 to 1985, the average domestic consumption of fat increased by 8.3 kg/capita, with 7.5 kg coming from visible sources. Approximately two-thirds of the visible fat available per capita in 1940 was from animal sources. The use of vegetable source fats has grown significantly, accounting for about two-thirds of the visible fats consumed in 1965, and 90 percent in 1985.[18]

Effects of Biotechnology

Compositions of oilseeds and their oils are being changed by classical plant breeding and biotechnology techniques. Examples include the development of canola, a low erucic acid–low glucosinolate seed by Canadian scientists in the 1960s and 1970s to replace traditional rapeseed, and sunflower and safflower seeds with high-oleic acid oils, which also are being grown. Other oilseeds in various stages of development include flaxseed and soybeans with reduced linolenic acid contents, soybeans with increased stearic acid content, and low-gossypol (glandless) cottonseed. Major investments have been made in oil palm biotechnology by PORIM—the Palm Oil Research Institute of Malaysia. Additionally, specialized enzyme systems are being developed for esterification/rearrangement of fatty acids on glycerol to produce fats similar to cocoa butter in composition.

EDIBLE FATS AND OILS EXTRACTION

Modern oilseed extraction plants are large, with capacities of up to 3000 to 4000 tons/day. Installations typically include facilities for unloading railroad cars, barges, or ships; storage of at least several weeks' supply of seed; solvent, meal, and crude oil storage tanks or facilities; processing facilities; and day facilities for personnel and supervisors. Additional space and facilities are required if refining and fats/oils processing facilities are done at the same location. A photo of an operating soybean extraction plant is shown in Fig. 8.3.

Oilseeds Solvent Extraction

Plant oils are derived mainly from: (1) dry, hull-covered, row crop seeds (soybeans, rapeseed/canola, cottonseed, peanuts, sunflower seed, and others); (2) the kernels of hard-shell nuts; and (3) pulpy flesh surrounding seeds (olive, palm, and avocado). Additionally, coconut flesh dried into copra, corn germ from dry or wet-milling processes, and bran from rice-polishing operations are rich in oil content and are handled somewhat like row crop oilseeds.

The oils of the ancient world were sesame and olive, probably because of their ease of extraction and high stability. Various societies have used plant oils for medicinal purposes. Later, they were used for lighting and then for food as extraction and refining techniques improved. Extraction technology has evolved from dispersing macerated or ground seed in hot water and skimming off the oil that rises to the surface (Cochin process), to pressing ground seeds (held in folded press cloths of human hair or other fibers) by lever or wedge presses, and later to mechanical and hydraulic batch presses. Some of these techniques are still used in remote areas of the world.

With growing demands for edible oils, industry adopted batch cage presses, and then continuous screw presses and solvent extractors. The latest technique, extraction with carbon dioxide in a high-pressure critical state, is being applied to high-value products such as essential oils, hops, and instant coffee. Usually, batch extractions are conducted because techniques for loading and unloading continuous extractors operating at 8000 to 12,000 psig have not been developed. More recently, however, batch extractions have evolved into "pulse-fed" continuous extractors operating at 8000 to 12,000 psig.

Oilseeds do not have fat cells like those of animals for storing fats. Instead, oil is stored in microscopic globules throughout the cells of dicotyledonous oilseeds, or in the seed germ or bran in monocotyledonous plants such as

Fig. 8.3. Photograph of an operating soybean extraction plant. (*By permission of Archer Daniels Midland Company, Decatur, IL.*)

maize (corn) and rice, respectively. Oil yields and processing costs very much depend on the effectiveness of preextraction operations to disrupt cells and free the oil for recovery. Classical soy processing involves denaturing protein in cell walls to enhance rupturing and disrupting of the cell contents. A generalized flow sheet for direct solvent extraction of many row crop oilseeds is shown in Fig. 8.4. Operations differ among various oilseeds, mainly in techniques of dehulling.

At the beginning of the 1980s, processing concepts were based on the classification of oilseeds into two groups. *High*-oil-content seeds were those containing over 30 percent oil on a dehulled or as-processed basis, and included rapeseed/canola, oil-type sunflower seed, peanuts, safflower seed, and copra. Typically, these seeds were dehulled (shredded in the case of copra), heated, flaked, and *hard screw-pressed* leaving 5 to 8 percent residual oil in the meal. Processors desiring to recover additional oil would first prepress high-oil seeds to 15 to 18 percent oil content using lighter-duty screw presses, and then solvent-extract the press cake to less than one percent residual oil content; this process is called *prepress-solvent extraction.* Hard pressing was considered too inefficient for a *low*-oil-content seed such as soybeans unless nearby markets were available for the oil-bearing meal. Thus, soybeans primarily have been *direct-solvent-extracted* to less than one percent residual oil content. Cottonseed was processed originally by hard press and later by prepress-solvent extraction, and now is processed mainly by direct solvent extraction techniques.

The introduction of the expander, a high-shear extruder with cut-flight screw, revolutionized oilseed extraction practices in the latter half of the 1980s. By 1990, over 80 percent of the soybean and cottonseed tonnage processed domestically used this technique. The expander heats, homogenizes, and shapes

Fig. 8.4. General flow sheet for extraction of field crop oilseeds.

seeds into porous collets (pellets) that are more dense (weigh more per unit volume) but are more rapidly extracted than flakes, approximately doubling throughput of continuous solvent extractors. Collets also drain more completely, greatly reducing steam costs for desolventizing the extracted meal. The Anderson International Company of Cleveland, Ohio, has patented an expander with a drainage cage (Fig. 8.5) to reduce the oil content of high-oil seeds to less than 30 percent, thus enabling the production of intact collets for direct solvent extraction from completely dehulled seeds such as sunflower seed and peanuts.

Introduction of the expander has enabled extraction plants to handle additional seed species, with the purchase of only minimal dehulling equipment where needed. Prepress-solvent-extraction facilities are rapidly being replaced by expander-direct solvent extraction processes, leaving two basic extraction processes in the modern oilseed industry—expander-direct solvent extraction, and hard press for applications where seed supplies are limited or other considerations do not warrant the construction of solvent extraction plants

Fig. 8.5. Anderson International Corp. Hivex-Series Expander (TM) model An-88-H with oil drainage section for preparing extraction collects from high-oil-content seeds. (*By permission of Anderson International Corp., Cleveland, OH.*)

or the expenses of skilled personnel and additional safety precautions for their operation. Hydraulic cage presses still are used in processing industrial crops such as castor seed, and for edible oils in developing countries.

Referring to Fig. 8.4, freshly harvested seed should be cleaned of trash, which harbors moisture that may accelerate the formation of free fatty acids (FFAs), and dried before storage. The maximum moisture content for holding seed for long periods without spoilage varies with storage temperature and between species, and generally is in the 8 to 12 percent range. If the seed has not been adequately cleaned before storage, this must be done before extraction processing to prevent clogging and damage of equipment.

Hulls are removed in a two-step process, called *dehulling* or *decortication*, in which they are first cracked and then separated by screening and aspiration to produce a processing intermediate material sometimes called *meats*. The major objective of removing the hulls is reduction of the fiber content in meals for feeding poultry and monogastric animals such as swine. However, the protein content also is increased, and the volume of material handled in the extractor is reduced. Earlier operations required partial retention of hulls in the meats to improve handling in flaking and screw pressing operations, but hulls are no longer required with the advent of expanders equipped with oil drainage cages. It is common practice either to leave sufficient hulls with the meats to just surpass minimum protein guarantees of meals, or to adjust high protein meals to industry trading standards by adding back hulls. Currently, dehulled (low fiber) soybean meal trades at 48 percent protein, soybean meal at 44 percent, cottonseed meal at 41 percent, and dehulled sunflower seed meal at 42 percent protein.

The meats are heated, with the addition of steam to increase the moisture content if needed. Heaters (cookers), desolventizer-toasters (DTs) and coolers used in the oilseeds processing industry often are of the shallow circular pan, sweep arm, multi-stack design shown in a later figure. In earlier techniques, seed was heated to about 74°C (165°F) before flaking to about 0.3 mm (0.012 in.) thickness. However, now it is realized that phospholipase, an enzyme that makes the phospholipids nonhydratable and more difficult to remove in water degumming, is highly active at this temperature, and seeds preferably are heated to less than 57°C (135°F) or to over 85°C (185°F) to avoid the range of maximum phospholipase activity. The seed then passes to the expander for rapid heating to 105 to 121°C (220–250°F), homogenization, and reshaping.

Prolonged hot-moist holding of flakes was used benefically before development of the expander; for example, (1) in the ALCON™ process to reduce production of nonhydratable phospholipids, trypsin inhibitors, and urease in soybeans, and (2) "cooking" of cottonseed to bind ("fix") gossypol, thus reducing its free levels in meals intended for poultry and monogastric animals. Because of the relatively short residence time of seed in the expander, a second pass, a subsequent hot-hold period or additional heat treatment in the DT after extraction, is used to deactivate anti-nutrients and to bind gossypol. Even though the expander homogenizes the seed, prior flaking enhances oil recovery but can be done at 0.5 mm (0.020 in.) thickness instead of the typical 0.3 mm (0.012 in.). The flakes or collets are cooled to about 6°C (10°F) below the boiling point of the solvent before entering the extractor.

Hardly any batch-type oilseed extractors remain, and modern solvent extractors mainly are of two basic designs. In *shallow bed-type* extractors, a 0.5- to 1-meter-thick layer of collets or flakes is pulled across a linear screen (Fig. 8.6), or conveyed on a woven belt, and repeatedly percolated with solvent. *Deep bed* extractors mainly are constructed as carousels with pie-shaped cells (*baskets*), that are alternately filled (2–3 meters deep), extracted, and unloaded. In some designs, the baskets rotate between the loading, extraction, drainage, and unloading stations (Fig. 8.7); in others, the baskets are stationary with the various stations revolving (Fig. 8.8). Solvent flow always is countercurrent to the direction of

Fig. 8.6. Cutaway drawing of Crown Model III Extractor™. Collects or flakes enter through the top left of the machine and are pulled in clockwise direction by a chain-link mechanism across the top extraction section and then the bottom section while percolated with miscella traveling in a countercurrent direction. (*By permission of Crown Ironworks, Minneapolis, MN.*)

Fig. 8.7. Cutaway drawing of Rotocell Solvent Extractor™. A carousel with deep-bed baskets rotates by a series of collet/flake loading, solvent extraction, drainage, and unloading stations. (*By permission of Davy-Dravo, Inc., Pittsburgh, PA.*)

Fig. 8.8. Cutaway drawing of French Stationary Basket Extractor™. The deepbed baskets are stationary while the collet/flake loading, extraction, drainage, and unloading stations revolve. (*By permission of French Oil Machinery Co., Piqua, OH.*)

the material being extracted, with the most oil-rich solvent (*miscella*) contacting the fresh collets, and reclaimed or virgin solvent used for the final rinse of the extracted material (*marc*). Continuous solvent extractors are built in various sizes up to 3000 to 4000/tpd (tons per day).

Many solvents have been proposed or used for extracting oilseeds, but later have been found ineffective in extracting oil or have been disallowed because of health concerns about residues in food and feed products and exposure of employees. Essentially all commercial oil extraction now is done with hexane, a petroleum refinery fraction with a boiling point of 65 to 68°C (149–155°F) that consists of 48 to 98 percent *n*-hexane with the balance being short-chain homologs and branched compounds. Currently, ethyl and isopropyl alcohol are attracting the most attention as alternative extraction solvents.[19]

The extracted, drained marc contains approximately 20 percent hexane solvent holdup for soybean collets and 33 percent for flakes, and is vaporized in a DT. Some DTs have cooling sections, but separate dryer-coolers (*DCs*) often are used in large installations. Steam and sometimes hot solvent are sparged into the marc as heat sources for volatilizing the solvent; however, the condensed steam must then be removed by drying. The toasting operation destroys enzymes and anti-growth factors such as trypsin inhibitor in soybeans, reduces meal protein solubility and microbial digestibility, and improves *rumen by-pass* or *rumen escape* in feeding cattle and sheep. Figure 8.9 shows a cutaway drawing of a Crown/Schumacher design desolventizer-toaster-dryer-cooler (*DTDC*).

In processing most oilseeds, hexane is stripped from the miscella by distillation to produce a crude oil. However, gossypol and other pigments become extremely difficult to remove if left warm in cottonseed oil for more than a few days. It is common practice for cottonseed oil mills either to send their crude oil immediately to an alkali refinery or to have a miscella or alkali refinery on-site to arrest the fixing of color.[20]

In miscella refining, miscella leaving the extractor contains 30 to 35 percent oil and is raised to approximately 65 percent oil by evaporation. The FFAs in the concentrate then are reacted with alkali (sodium hydroxide solution) to produce soaps that are removed by centrifugation. Then the solvent is removed from the refined miscella by evaporation, and the soapstock is placed on the meal in the DT to recover its solvent. Hexane vapors from the miscella and the DT are condensed, and the solvent is recycled to the extractor for reuse. Noncondensable gases are passed through a mineral oil stripper to recover the last traces of hexane. A well-designed extraction plant should lose no more than about 0.5 gallon of solvent per short ton (2000 lb) of oilseed extracted.

Screw Pressing

The objective in screw pressing is to preheat (cook and dry) the seed tissue to the point

Fig. 8.9. Cutaway drawing of Crown/Schumacher desolventizer-toaster-dryer-cooler (DTDC). Note the stacked design of tray (ring) heaters-coolers with sweep arms, and steam-jacketed and perforated bottoms. Stack cookers generally have solid steam-jacketed bottoms for heating oilseeds. (*By permission of Crown Iron Works Company, Minneapolis, MN.*)

that the cell walls are brittle and rupture readily upon pressing. The seed generally is dehulled, tempered, flaked, and cooked to low moisture (2–4 percent) levels before hard pressing (Fig. 8.4). Concerns have been expressed about destruction of the essential amino acid lysine and reduction of nutritional quality and economic value of protein meals made by high-temperature, hard-press processes. Pressing typically is done in a two-stage continuous screw press or expeller, as shown in Fig. 8.10.

Extraction of Olive, Coconut, Palm and Palm Kernel Oils

Oil is extracted from the fruit of the olive tree, *Olea europea*, by milling to remove the pulpy flesh from the hard seed. The resulting paste then is: (1) pressed in hydraulic presses that use disk filters ("capachestas"); (2) diluted with water, homogenized, and allowed to settle by gravity; or (3) separated by using a continuous horizontal decanter (Fig. 8.11), sometimes followed by a vertical disc-type centrifuge for polishing (Fig. 8.12).[21]

Coconuts, *Cocos nucifera*, are hand-cut from the coconut palm with knives attached to long poles and dehusked, and the nut is cracked to drain away the coconut water. Then, the flesh is separated from the adhering shell and dried, raising the oil content from approximately 30 percent in the meat to 50 to 65 percent in the resulting copra. Considerable amounts of copra are prepared by hand and sun-dried to supplement family income in tropical countries. However, centralized coconut mechanical husking and dehulling facilities—which dry coconut meat using heat generated from burning the husks and offer the advantages of improved moisture control and reduced mold spoilage and aflatoxin problems—are growing in numbers. In the oil extraction process, copra is: (1) cleaned by shaker screens to remove trash; (2) size-reduced by hammer or attrition mills,

Fig. 8.10. Anderson Expeller™ Press, model 55″ TMSD. Note elevated conditioner carrying product from left to right across top of press; vertical screw press section (right side); second (higher pressure) screw press section carrying product from right to left. Other brands of hard presses have both stages on one shaft. (*By permission of Anderson International Corp., Cleveland, OH*).

or fluted roller mills; (3) optionally flaked; (4) additionally dried if necessary; and (5) hard-pressed or prepress-solvent-extracted.

Two distinctly different types of oils are produced from the fruit of the Southeast Asia and African oil palm, *Elaeis guineensis*, and its hybrids with the South and Central American palm *E. oleifera*. Palm oil is otained from the fleshy part of the fruit, which resembles an oversized olive about the size of a small chicken egg. Palm kernel oil is derived from the kernel within the nut. Well over 98 percent of the fatty acids in palm oil belong to the C:16 and C:18 group, whereas approximately 64 percent of the fatty acids in palm kernel oil consist of the C:12 and C:14 lauric group.

A palm tree produces 10 to 15 fresh fruit bunches throughout the year, weighing 5 to 23 kg (10–50 lb) each. The bunches are cut from the tree with knives attached to long poles and are transported to the oil mill. There, they are sterilized by steam at about 40 psig for 50 to 75 minutes to deactivate lipase enzymes and loosen the fruits from the stalk. The fruits are knocked loose from the stalk in thresher drums, and passed through a digester to convert the fleshy pulp to mash. Then the mash is pressed by twin-screw expellers or hydraulically to yield red crude oil. The shells of the nuts are cracked, and the kernels are separated, dried, and bagged for later solvent extraction in a fashion similar to the processing of row crop oilseeds.

Extraction of Lards, Tallows and Fish Oils

Dry and wet rendering methods are used for extraction of animal fats and fish oils. In dry rendering of animal fats, the fatty tissue is

1 Conveyor screw drive
2 Bowl drive
3 Bowl bearings
4 Cyclo gear
5 Distributor
6 Conveyor screw
7 Centrifugation space
8 Bowl
9 Regulating plate
10 Feed
11 Discharge of the clarified liquid
12 Frame
13 Solids discharge
14 Drive motor
15 Clutch

Fig. 8.11. Schematic drawing of Westfalia Model CA 450 Continuous Clarifier Decanter. (*By permission of Centrico, Inc., Northvale, NJ.*)

ground and heated to drive off moisture, a process that makes the cell walls brittle by denaturing the protein. The resulting slurry of oil and cracklings is passed through a screw press to expel the oil, and the press cake then is ground into meat or meat by-product meal. Batch-type cookers of various types mainly are used, followed by continuous single-stage presses. Where batch-type cage presses still are used, the cake is formed in the shape of large wheels, sometimes called *greaves*.

In wet rendering of animal products, the flesh is ground and heated to denature the protein and release moisture and fat. Continuous decanters then are used to remove the liquid fraction, and the residue is frozen and sold as *partially defatted beef or pork tissue* for use in processed meats. The liquid portion is additionally heated to help break the emulsion, and the fat is separated from the aqueous phase by continuous centrifuges.

Much of the fish that is processed is cooked in batch or continuous cookers to coagulate the protein, and passed through a continuous screw press to remove the liquid portion. The liquid then is separated by continuous decanters

1 Timing unit
2 Heavy phase outlet
3 Feed
4 Light phase outlet
5 Flush water inlet
6 Pressure switch
7 Heavy phase centripetal pump
8 Light phase centripetal pump
9 Distributor
10 Disc set
11 Sediment holding space
12 Sediment ejection ports
13 Piston
14 Closing water
15 Opening water
16 Closing chamber
17 Bowl valve

Fig. 8.12. Schematic drawing of Westfalia Model RSA 150 Refining Separator with self-cleaning bowl. (*By permission of Centrico, Inc., Northvale, NJ.*)

into aqueous and oil fractions, and the latter then is passed through a continuous polishing centrifuge. The solid residue is dried and ground to produce fish meal.

Essential Oils

Essential oils are extracted in several ways: with fat, by distillation, and by batch solvent extraction. The oldest method, used today only for extremely valuable essential oils, is *enfleurage-défleurage*. This consists of layering blossoms of the material to be extracted between fat-coated glass plates and allowing the essential oil to perfuse into the fat. The blossoms are renewed daily. At the end of the season, the fat is scraped from the plates, melted, poured into containers, and sold as *pomade*, or is batch-extracted with cold ethanol and sold as *extraits*. Another technique consists of macerating the material and extracting with hot fat. The most common extraction process for essential oils is steam distillation of a mash of the leaves or seeds to be extracted. Selective solvent batch extraction also is used.

EDIBLE FATS AND OILS: REFINING, PROCESSING, AND COMPOUNDING

The objectives of refining and processing fats and oils include: (1) removal of free fatty acids, phospholipids (gums), color and off-flavor/odor compounds, and toxic substances to produce light-colored, bland products with long shelf lives; (2) obtaining a mixture of triacylglycerols with the desired solid content profiles over the range of product use; and (3) preparation and storage of semisolid products with desired textures.

Solid Fat Content

Crystallization characteristics of fats and oils were described in the section on polymorphism. The size, abundance, and stability of crystals are important in food processing/food service ingredients and consumer products. Clear salad oils are preferred, and are made by selecting low-melting-point oils that are *winterized* by chilling to crystalline the higher-melting triacylglycerols, which are then removed by filtration.

Margarines, shortenings, and coating fats are mixtures of fat crystals in oil, whose characteristics are described and controlled by Solid Fat Content (SFC) or Solid Fat Index (SFI) techniques. The results of the two analytical techniques are similar but not identical.

SFI curves of butter, stick and tub (soft) margarine, and general-purpose shortening oils are shown in Fig. 8.13. Note that stick margarine can be formulated to be softer (contain less solids) than butter when removed from the refrigerator, and to remain firmer at warm room temperatures. Tub margarines can be made soft and spreadable when taken directly from the refrigerator, with relatively small change in texture over a wide range of temperatures.

Fig. 8.13. Solid Fat Index curves of butter, stick and tub margarine fats, and a general-purpose shortening.

Fig. 8.14. Solid Fat Index curves of cocoa butter, cocoa butter equivalent (CBE, used as an extender), and higher-temperature-melting cocoa butter substitutes (CBSs).

Fats must melt completely just below body temperature (37°C) to avoid coating the roof of the mouth and giving the product a greasy feeling. This explains why the shortening depicted in Fig. 8.12 would taste very greasy, and why some stick margarines are slightly greasy. Another important concept in semisolid fat product technology is the *plastic range*, or the temperature span through which a product remains soft and workable but without being fluid; 17 to 25 SFI has often been used as the plastic range for shortenings.[22]

Figure 8.14 presents SFI curves of cocoa butter (from the kernel of *Theobroma cacao*), a cocoa butter equivalent (CBE) used as a cocoa butter extender, and two cocoa butter substitutes (CBS) or replacers. Cocoa butter, the main component of chocolate, has a pleasant mouthfeel because it completely melts from a solid product into a liquid in a short temperature span. Considerable energy uptake is required, and the melting of chocolate has an agreeable cooling effect in the mouth. CBEs are made by compositing triacylglycerols containing primarily C16 and C18 fatty acids. They are used as extenders, as enhancers to slightly raise the melting point of cocoa butter, and as replacements. CBEs are compatible with cocoa butter, and can extend marketing opportunities of confections to hot weather seasons and warmer climates. CBSs consist primarily of lauric acid–type fats (primarily C12 and C14 fatty acids from palm kernel and coconut oils). Although CBSs can simulate the characteristics of cocoa butter, these substances are not compatible and are used independently.

The production of semisolid oil and fat products with specific desired textures includes the selection of fat/oil mixtures with the desired SFI/SFC profiles and working the product into a stable crystal state. Thermodynamic principles dictate that fats eventually will undergo polymorphic changes to reach the most stable β state. However, changes can be appreciably delayed kinetically by: (1) selecting mixtures of triacylglycerols that co-crystallize as small crystals; (2) specifically including oils that are β'-formers (such as corn, cottonseed, fish, linseed, olive, palm, and peanut oils and tallow) in preference to β-formers (such as soybean, canola, sesame, and sunflower oils and lard)[23]; (3) including emulsifiers that interrupt the growth of crystals; (4) chilling-working the product into specific polymorphic states.

Regarding the last point, many melted shortenings normally would cool into mixtures

Fig. 8.15. Cut-away drawing of a Votator™ scrape-surface heat exchange unit used for making plasticized shortening. (*By permission of Cherry-Burrell Process Equipment, A United Dominion Company, Louisville, KY.*)

of crystals in free oil, which are difficult to handle and have limited air-entrapment abilities in making cakes and pastries. Chilling the same stock to form small crystals and aeration with inert gases in scrape-surface heat exchangers called *Votators*™ (Fig. 8.15) produce smooth-textured, semisoft products that are easy to handle.

Up to six polymorphic forms of cocoa butter crystals have been reported. It is common practice to work chocolate into very stable β' or small β crystal forms by slow agitation, careful control of temperatures, and tempering (holding) freshly cast products under specific conditions before release to the trade. *Bloom*, a grayish defect, occurs when chocolate or other types of coating fats have been allowed to partially melt, and some oils rise above the sugar and coloring matter in the product and recrystallize.

Degumming and Refining

Crude fats and oils are processed by the general scheme shown in Fig. 8.16, with modifications or exceptions for specific species. The phospholipids (lecithins) must be removed to avoid darkening of the oil during high-temperature deodorization and in deep-fat-frying applications. This removal typically is accomplished during the alkali refining process or in a separate water/acidic water

Fig. 8.16. Generalized flow sheet for refining and processing fats and oils.

degumming step before alkali refining. Crude soybean oil has an unusually high (2–3.5 percent) phospholipid content among oils, and often is degummed in a separate operation to no more than a 300 ppm level (as phosphorus) to avoid precipitation during shipping and storage. Levels of 10 ppm or less of phosphorus are achieved in the better-quality refined soybean oils. Degumming is achieved by mixing crude soybean oil with water to hydrate the phospholipids and enable their removal by centrifuge. Citric and other acids sometimes are added, in a step called *super degumming*, to help remove phospholipids that are not hydrated by water alone. The gums then may be purified into commercial lecithins, or placed back on the meal in the DT if a market is not available.[24]

Degummed soybean oils, or crude oils of other species, next are reacted (*neutralized*) with sodium hydroxide solution to form sodium salts of the fatty acids, which are removed as soapstock by a continuous centrifuge.[25] The soapstock also occludes remaining phospholipids, some color and flavor compounds, and several percentage points of neutral (salable) oil for each percentage of FFAs present. The soapstock can be put back on meal in the DT for drying if refining is done adjacent to an extraction plant, or acidified to again form fatty acids and sold to the oleochemicals industry. The oil then is water-washed and centrifuged one or two times to remove residual soaps.

Bleaching and Deodorization

Next, the refined oil is bleached (reacted with activated clays, diatomaceous earths, or recoverable silicas) to adsorb phospholipids, color pigments, soaps, peroxides, aldehydes, and other polar compounds, and is recovered by filtration. Spent bleaching earths in the filter cake contain pigments, occasional residual pesticides, aldehydes, and other compounds removed from the oil, and they often present disposal problems.[26]

Deodorization consists of heating the oil to high temperatures under high vacuum with the addition of stripping steam to remove remaining odor-causing compounds, volatile pesticides, products of oxidative rancidity such as peroxides, aldehydes, tocopherol-like natural antioxidants, and residual amounts of free fatty acids.[27] The deodorizer condensate may be collected and processed to recover natural antioxidants. The final refined, bleached, and deodorized product is called *RBD oil.*

As an alternative, the degummed oil may be subjected to minor bleaching and then physically refined, using a high-temperature/high-vacuum process that strips away the compounds normally removed in the deodorizer plus the fatty acids.[28] The major advantage of physical refining is elimination of neutral oil losses that occur by occlusion in both miscella and alkali refining. Palm oil has low phospholipid content and generally is physically refined, but row crop oils need to be degummed before physical refining can be considered.

Hydrogenation, Rearrangement, and Thermal Fractionation

These three processes are mainly concerned with obtaining fats and oils with desired stability and solid content properties. In theory, they may be used alternatively or in any sequence; but, in practice, have been applied more to certain species of fats and oils than to others.

Hydrogenation is the process of adding hydrogen with the assistance of a catalyst (often nickel) to an unsaturated (double) bond of a triacylglycerol fatty acid:

However, a polyunsaturated hydrocarbon chain can isomerize after hydrogenation to form either *cis*- or *trans*-isomers at the remaining unsaturated sites. Unsaturated oils in nature normally are in the *cis*-form, and concerns about potential roles of hydrogenation-produced *trans*-forms in circulatory and heart diseases have been voiced.[29] The ratios of *cis*:*trans* forms produced can be

partially controlled by the catalyst and the temperature–pressure conditions selected for hydrogenation. With continued hydrogenation, polyunsaturated fats can be changed to fully saturated fats with high melting points. However, hydrogenation typically is used with discretion to achieve desired intermediate forms. "Brush" or light hydrogenation often is used to reduce the linolenic acid content and increase flavor stability of bottled soybean oil. Also, arrays of fats with different melting points are prepared as blending stocks for making shortenings, margarines, and specialty fats.[23]

Hydrogenation is applied broadly to row crop oilseed oils.

cis-9, cis-12	trans-10, cis-12	trans-10 cis-12 trans-11 cis-11
	cis-9, trans-11	cis-9 trans-11 cis-10 trans-10
Linoleate	Conjugated dienes	Monoenes

Rearrangement, also called *interesterification*, is a means of changing the structure of a triacylglycerol. By controlling purity of the triacylglycerol stock, it is possible to rearrange the location of fatty acids on the glycerol, move (*transesterify*) fatty acids from one glycerol to another, or displace a fatty acid on a glycerol with a fatty acid introduced to the mixture.[30]

Fatty acids in natural fats are not randomized, but occur in specific distributions; however chemical interesterification randomizes both their distributions between triacylglycerols and locations on the glycerol molecule. For example, the natural distribution of saturated and mono- and polyunsaturated fatty acids in cocoa butter is: SSS—2.2 percent, SSU—84.6 percent, SUU—11.1 percent, and UUU—0.5 percent, with 1.6 percent unknown; whereas the distribution in randomized cocoa butter is: SSS—23.9 percent, SSU—43.8 percent, SUU—26.8 percent, and UUU—5.5 percent.[31] Depending on the fat, rearrangement can improve or damage the desirable SFC, plastic range, and melting characteristic. The functionality of lard often is improved by transesterification, but the highly desirable sharp melting characteristic of cocoa butter is destroyed by rearrangement. Rearrangement has been used very effectively to improve the functionality of lard.

Thermal fractionation is the process of chilling an oil to precipitate crystals of the higher-melting triacylglycerols, and removing these crystals with continuous filters.[32] *Dewaxing* is the process of chilling the oil and removing usually low levels of waxes by centrifugal force with the help of a wetting agent, or by filtration with the aid of filter aids. *Winterization* is the process of reducing the cold test of an oil by chilling followed by removal of the higher-melting triacylglycerols. *Dry fractionation* is the process of separating a fat or oil into two or more fractions by a

$$\begin{array}{l} CH_2OOC\text{–}R_1 \\ | \\ CHOOC\text{–}R_2 \\ | \\ CH_2OOC\text{–}R_3 \end{array} + R_4\text{–}COOH = \begin{array}{l} CH_2OOC\text{–}R_4 \\ | \\ CHOOC\text{–}R_2 \\ | \\ CH_2OOC\text{–}R_3 \end{array} + R_1COOH$$

Triacylglycerol Fatty acid New triacylglycerol Fatty acid

$$\begin{array}{l} CH_2OOC\text{–}R_1 \\ | \\ CHOOC\text{–}R_2 \\ | \\ CH_2OH \end{array} + R_3\text{–}COOH = \begin{array}{l} CH_2OOC\text{–}R_1 \\ | \\ CHOOC\text{–}R_2 \\ | \\ CH_2OOC\text{–}R_3 \end{array} + H_2O$$

Diacylglycerol Fatty acid Triacylglycerol

combination of chilling, seeding with nuclei if necessary, and filtration. The liquid (oil) fraction is called *olein* and the solid fraction *stearin*. Dewaxing, winterization, and fractionation, in the presence of a solvent, are called *wet* or *miscella fractionation*.

Fat and Oil Products

Salad and *household cooking oils* typically are winterized before bottling. However, some premium-priced, cold-pressed, filtered low-fatty-acid oils are sold in health food outlets with the claim that they have received no chemical treatment.

Liquid shortenings consist of selected oils to which a crystal inhibitor (fatty acid, solid triacylglycerol, lecithin, glyceryl or sorbitan mono- or diester, or other compound) has been added; the product is chilled to retain the oil in the α-crystal form. Liquid shortenings are available in cloudy or clear form.

In making *shortenings*, oils and hardstocks made previously by hydrogenation are blended to obtain the desired plastic range, and are chilled and aerated using a Votator™ or other scrape-surface heat exchanger. In making *margarines*, water, milk solids, flavorings, and color are blended to obtain a mixture containing 80 percent fat and then are chilled by a scrape-surface or other type of heat exchanger. Time must be allowed during the process, or in a later tempering room, for the desired crystal forms to materialize before these products are shipped.

Notes on Selected Species

Soybeans (*Glycine max*). Soybeans originated in China, and were cultivated as early as 2836 B.C. The oil of this seed has presented numerous processing challenges because of its high phospholipid content and a tendency to "revert" to a beany flavor in the early days of processing technology. Recommendations for processing include exclusion of copper and iron in equipment and in lines, avoidance of light, and storage under a nitrogen blanket.

Cotton (*Gossypium arboreum* and *G. herbaceum*). Cotton was grown in the Indus River Valley (modern Pakistan) as early as 3000 B.C., "New World" cottons were grown in Peru (*G. hirsutum*) in 2500 B.C., and Sea Island cottons (*G. barbadense*) were grown in the Caribbean Islands. Most of the world's cotton today is *G. hirsutum*, with fuzzy seed that requires a delintering saw before dehulling and subsequent processing. Pima and Egyptian cottons (*G. barbadense-type*) have "naked" seed and can be processed directly. Cottonseed is among the first examples of a reclaimed by-product in our industrial world. Invention of the cotton gin by Eli Whitney in 1793 led to increased domestic production of cotton and also to seed disposal problems. The state of Mississippi passed the first anti-pollution law in the nation in 1857, prohibiting throwing cottonseed into rivers and requiring its orderly disposal from gins located less than one-half mile from a town.

Cottonseed contains the polyphenolic green-yellow pigment gossypol, 1,1′,6,6′,7,7′-hexahydroxy - 5,5′ - diisopropyl - 3,3′ - dimethyl(2,2′-binapthalene)-8,8′-dicarboxaldehyde, in discrete bodies called *glands*. Processing challenges include removal of gossypol, or its 15 or more derivatives, from the oil to reduce color, and deactivating its toxic effects (through binding to the protein with moist heat) to enable feeding the meal to monogastric animals.

Cottonseed oil also contains the cyclopropenoic acids, C18:CE malvalic acid (8,9-methylene-8-heptadecenoic acid) and C19:CE sterculic acid (9,10-methylene-9-octadecenoic acid), which form the pink color in the Haphen test used to detect its adulteration of higher-priced olive oil.

Gossypol

Sterculic acid

Malvalic acid

Rapeseed/Canola. These belong to the turnip rutabaga, cabbage, Brussels sprouts, and mustard family of crops that can be grown at low temperatures and moderate humidity. Three species have been grown as oilseeds: *Brassica napus*, known in Europe as rape, oil rape, swede rape, and Argentine rape; *B. campestris*, known as rapeseed, oil turnip, turnip rape, and Polish rape; and *B. juncea*, known as leaf mustard, brown mustard, Oriental mustard, and Indian mustard.[33] *B. campestris* was grown in India as early as 2000 to 1500 B.C.

This crop is the classic example of an effective crop improvement program mobilized by one country in response to the impending loss of international markets. When it was reported in the early 1960s that the erucic acid in rapeseed oil can cause heart damage and other diseases, the Canadian government established a crash development program that led to release of the first low-erucic-acid rapeseed (*LEAR*) in 1969. Continuing efforts to develop varieties with even lower erucic acid led to the later release of "*double zero*" (low-erucic, low-glucosinolate) varieties, which were named *canola* in 1980. Because of reduced glucosinolate levels, considerably more of the meal could be used in animal feed than earlier. Canola oil, which contains less than 2 percent erucic acid compared to 20 to 40 percent in earlier rapeseeds, was granted GRAS (generally recognized as safe) status by the United States Food and Drug Administration in 1985.

However, introduction of canola did not solve all problems. An industrial market for erucic acid had developed earlier. High-erucic-acid rapeseed then was imported from Northern Europe for extraction, followed by efforts to develop industrial rapeseeds with increased erucic acid content as well as development of *crambe* (*Crambe abyssinica*) specifically for its erucic acid. At the current state of development, equipment corrosion and poisoning of hydrogenation catalysts by sulfur in canola oil still are problems. Also, oil from the new canola varieties favors formation of β-crystals, in contrast to β'-crystal formation in the earlier rapeseed varieties.

Sunflower (*Helianthus annus* var. *Marcocarpus*). This is a New World crop, known to have been grown in Arizona–New Mexico in 3000 B.C. and in the Mississippi–Missouri Basin at least since 900 B.C. The crop was taken to Europe by early explorers and was developed as a source of edible oil in Russia. The gray and white-striped seed is known as the *confectionary-type*. Its kernel is loose within the achene, and the seed usually is shelled and eaten. No major antinutritional factors are known to exist in the raw seed. *Oil-type* varieties predominantly are black-hulled, with seeds about one-third the weight of confectionary-types. The seed adheres tightly to the hull, and provides a challenge in dehulling. It has been repeatedly shown that, although the setting of seed by sunflower plants is not as sensitive to day length as that by soybeans, the oleic acid content is inversely related to the temperature of seed maturation. Thus, users desiring polyunsaturated oil with high linoleic acid content have specified northern-grown sunflower oil, whereas those desiring extended oil stability, as in production of fried foods, have preferred oil from seeds matured during the summer in the southern states. Sunflower varieties with high-oleic-acid content (85–92 percent) have been developed within the last decade, with intentions of using the oil as a feedstock in chemicals manufacturing processes. Also, high-oleic confectionary-type sunflower seed has been developed, enabling extension of the shelf-life of roasted seed.

Peanuts, Ground Nuts (*Arachis hypogae*). Also a New World crop, peanuts are known to have been grown in the Upper Plata Basin of Bolivia in 2000 B.C. They were taken to Europe by early explorers, and were returned

to the southeastern United States from Africa by slave traders. Broad cultivation did not occur domestically until the early 1920s. Most of the peanuts grown in the world are processed into oil, but most domestic-grown peanuts are consumed as food, with over one-half of the crop production used in making peanut butter.

NONIONIC SURFACTANTS—EMULSIFIERS

Nonionic surfactants and emulsifiers are compounds whose molecules have both aqueous (polar) and alkane (nonpolar) compatible sectors. Their molecules have regions that are sufficiently similar to become part of either system, with the other sectors sufficiently dissimilar to interrupt continuation of that system. For example, when added to a crystallizing fat, some nonionic surfactants may orient themselves to become part of the crystal, and will prevent further replication thus limiting crystal size. Likewise, some will react wth gelatinized starch in aqueous systems and prevent its recrystallization (retrogradation). Further, surfactants may orient around discrete droplets to stabilize water in oil (W/O) or oil in water (O/W) emulsions. When this occurs, the compatible end becomes associated with the discrete droplet, leaving the other end turned outward to associate with the compatible continuous phase.

The major groups of commercial emulsion stabilizers include: (1) glycerol esters, favoring W/O emulsions; (2) esters of monoacylglycerols with hydroxycarboxylic acids (including lactic, succinic, malic, and tartaric); (3) sodium stearoyl-2-lactylate (SSL), favoring O/W emulsions; (4) fatty acid monoesters of ethylene glycol; (5) sorbitan fatty acid esters known as SPANS (W/O) and TWEENS (O/W); (6) phospholipids; (7) water-soluble gums, including gum arabic, tragacanth, xanthin, agar, pectin carrageenan, and methyl- and carboxymethylcellulose; and (8) proteins.

Glycerol has three exposed hydroxyl groups, resulting in a compound completely miscibile in water at all concentrations. Fatty acids have both hydrophilic (water-attracting) and lipophilic (oil-attracting) ends. As they are esterified to glycerol, a molecular structure that originally was primarily water-soluble becomes increasingly oil-soluble. Finally, when all three positions are esterified with non-hydroxy fatty acids, the molecule is nonpolar and soluble only in organic solvents.

Limited amounts of monoacyglycerols and diacylglycerols are present in extracted crude oils. Additional quantities can be made by chemical or enzymatic hydrolysis, or by esterification of glycerol with one or more fatty acids. Also, phospholipids normally are present in extracted plant crude oils and have at least one chain with a water-soluble unattached end. At best, fatty acids, monoacylglycerols, and diacylglycerols are W/O emulsifiers. Additional types of surfactants can be made by esterifying fatty acids to compounds other than glycerol, including propylene glycol, surcrose, and sorbitol (resulting in SPANS™). TWEENs™, which are O/W emulsifiers, detergents, and solubilizers, are made by adding polyoxyethylene (POE) chains to SPANS.

The hydrophilicity of nonionic surfactants can be characterized numerically as their hydrophile–lipophile balance (HLB). An HLB value of 3 to 6 indicates that the compound is a likely W/O emulsifier; 7 to 9, a wetting agent; 8 to 13, an O/W emulsifier; 13 to 15, a detergent; and 15 to 18, a solubilizer (of oil or other nonpolar compounds into water).[34] HLB values of some common compounds are presented in Table 8.7.

A *phospholipid* is defined as any lipid containing phosphoric acid, and a *glycerophospholipid* is any derivative of glycerophosphoric acid that contains at least one O-acyl, O-alkyl-l-enyl, or O-l-alkyl group attached to the glycerol residue. The terms phosphatide and phosphoglyceride are no longer recommended. Under stereospecific nomenclature, the phosphate is assigned to the *sn*-3 position. By this convention, phosphatidylcholine (PtdCho), a major component of soya lecithin, is called 1,2-diacyl-*sn*-glycero-3-phosphocholine.[35] General structure formulas for the more common phospholipids

in plant oils are shown below:

R = Fatty acid

$R'' = H$	Phosphatidic acid
$R'' = C_6H_6(OH)_6$	Phosphatidylinositol
$R'' = CH_2CH_2NH_3^+$	Phosphatidylethanolamine
$R'' = CH_2CH_2N^+(CH_3)_3$	Phosphatidylcholine
$R'' = CH_2CH(NH_3^+)COOH$	Phosphatidylserine

Common phospholipids in plants

TABLE 8.7 HLB (Hydrophile–Lipophile Balance) Numbers for Some Surfactants

Name	*CAS Number*	*HLB Value*
Oleic acid	112-80-1	1.0
Acetylated monoglycerides (film formers)	—	1.5
Sorbitan trioleate (SPAN 85[1])	26266-58-6	1.8
Glycerol dioleate	25637-84-7	1.8
Sorbitan tristearate (SPAN 65[1])	26658-19-5	2.1
Glycerol monooleate	25496-72-4	3.4
Glycerol monostearate	31566-31-1	3.8
Sorbitan monooleate (SPAN 80[1])	1333-68-2	4.3
Sorbitan monostearate (SPAN 60[1])	1338-41-6	4.7
Soy lecithin	8020-84-6	8.0
Sodium stearoyllactylate (anionic type)	18200-72-1	8.3
POE sorbitan monooleate (TWEEN 81[1])	9005-65-6	10.0
POE sorbitan monostearate (TWEEN 60[1])	9005-67-8	14.9
POE sorbitan monooleate (TWEEN 80[1])	9005-65-6	15.0
POE stearic acid (monoester) (TWEEN 20[1])	9004-99-3	16.9

[1]Brand names by Atlas.

An HLB value of 8.0 is shown in Table 8.7 for lecithin, but manufacturers are able to supply lecithins with values of 2 to 12. The reader is cautioned that the terms "phosphatidylcholine" and "lecithin" sometimes are used interchangeably, especially in nutrition literature.

HEALTH AND CONSUMER INTERESTS

Non/Low-Caloric Fat Substitutes

Although various health groups recommend that no more than 30 percent of calories be consumed as fats and oils, the national

average is approximately 37 percent. Research on military and astronaut rations has focused on increasing the caloric density of food to save both weight and cube (volume) without inducing nutritional problems. However, the most publicized research and development has been on ingredients that impart the desirable characteristics of fat at reduced calorie levels. Proposed replacements have been protein-based, carbohydrate-based, or synthetic (fatty acids and long-chain alcohols in structures that are resistant to hydrolysis by digestive enzymes).[36,37]

Protein-based substitutes are derived from egg and milk proteins that have been heated and blended in a process called microparticulation, which shears and shapes the heat-coagulated protein gels into small spheroidal particles that are below the perceptual threshold of the tongue, thus giving a rich and creamy sensation in the mouth. The only product granted GRAS status for use in frozen desserts by the FDA thus far is Simplesse[R], a product made by The NutraSweet Company, Skokie, IL. This is a hydrated protein, containing about 1.3 kcal/g, that is used to replace fat (9 kcal/g) in frozen desserts. Other potential applications include dairy products, salad dressings, mayonnaise, and margarine. Kraft General Foods, Glenview, IL, has petitioned for approval of Trailblazer[R], a product reported to be similar but made by a different process.

Carbohydrate-based replacements rely on a viscosity increase and smooth gel-like textures to simulate the properties of fats and oils. These substitutes include: gums, hydrophilic hydrocolloids that increase product viscosity and improve emulsion stability; polydextrose, a polymer of dextrose with small amounts of sorbitol and citric acid made by Pfizer Chemical Division, New York; and a variety of corn, tapioca and potato starch maltodextrins made by various starch processors. Neither the protein- nor the carbohydrate-based replacements can be used as frying or dry coating oils.

Synthetic compounds are fatlike materials that include fatty acids or long-chain alcohols esterified in structures that are not hydrolyzed by enzyme activity. None of these compounds has been approved yet by the FDA.

The best-known material under consideration is Olestra, a proposed generic name for sucrose polyester developed by the Procter & Gamble Company, Cincinnati, OH. This compound contains six to eight long-chain fatty acids, primarily C12 to C24, esterified to sucrose. It is made by converting RBD oil into methyl esters by reacting the triacylglycerol with methanol using sodium or potassium methoxide as the catalyst, and then combining the washed esters with sucrose using an alkaline catalyst. To best simulate their respective properties, Olestras preferably are made from fatty acids of the oils that are to be replaced, such as cottonseed or corn oil. Olestras are intended for entirely replacing fats and oils in deep-fat frying operations, coating snack foods, and other applications. It also is claimed that the problem of anal leakage (spotting of clothes by nonassimilated fats that pass intact through the digestive tract) has been solved by the development of Olestras that do not separate because their consistency is similar to that of the feces.

Other proposed synthetic compounds include esterified propoxylated glycerols (EPG), developed by ARCO Chemical Company, Newton Square, PA. The product is prepared by reacting glycerine with propylene oxide to form a polyether polyol that is esterified with one or more fatty acids. Frito-Lay, Inc., Dallas, TX, has developed DDM (dialkyl dihexadecymalonate), a fatty ester of malonic and alkylmalonic acids for high-temperature applications. Use of this product is reported to result in a 33 percent reduction in calories and a 60 percent reduction in fat content. The Best Foods Division of CPC International, Inc., Englewood Cliffs, NJ, is evaluating TATCA (trialkoxytricarballate), tricarballic acid esterified with fatty alcohols.

Nutrition and Health

Interest in the role of dietary fats in health has grown rapidly in the last four decades, focusing first on essential fatty acids and then on the role of fats and associated compounds

in cardiovascular diseases. This is still a controversial and rapidly changing field, and the reader is urged to seek the latest information in the appropriate technical journals.

Currently, linoleic acid (C18:2*n*6) is considered a dietary essential fatty acid, as is arachidonic (C20:4*n*6)—but only when a linoleic acid deficiency exists. α-Linolenic acid (C18:3*n*3) generally is considered essential, but it still is undergoing review. Also, *n*-6 and *n*-3 polyunsaturated fatty acids have been shown to be precursors of eicosanoids (prostaglandins, thromboxanes, and leukotrienes), which are important in the regulation of many physiological processes. Eicosapentaenoic acid (EPA, C20:5*n*3) and docosahexaenoic acid (C22:6*n*3) are the major polyunsaturated fatty acids in fish, and consumption of fish rich in EPA is thought to modify blood platelet function and inflammatory responses. These types of interest have resulted in increased self-dosage with fish and other omega-3 oil sources by the general public. A linoleic acid intake to 1 to 2 percent of total dietary calories is considered sufficient to prevent biochemical and clinical evidence of deficiencies in human adults, and the American Academy of Pediatrics has recommended that infant formulas provide at least 2.7 percent of their energy as linoleic acid. The minimal daily intake of linoleic acid for adults is considered to be more than met by diets in the United States, and it is cautioned that large amounts of polyunsaturated fatty acids may increase the need for vitamin E, a fat-soluble, antioxidant nutrient. The NRC recommends that individual intakes of polyunsaturated fatty acids and of saturated fats each not exceed 10 percent of caloric intake.[2] However, because essential fatty acid deficiency has been observed exclusively in patients with medical problems of fat intake or absorption, a rationale may well exist for modified intakes supervised by physicians.

INDUSTRIAL FATS AND OILS

Fats and oils for industrial use typically are handled similarly to edible products, but usually with lower sanitation requirements.

Inedible Rendered Fats and Oils

Animal by-products, condemned carcasses and parts from packing plants, trimmings and bones from butcher shops, spent frying fats/oils, and fats collected from restaurant and food service kitchen grease traps are dry-rendered to produce inedible fats and oils.

Types of Industrial Fats and Oils Sought

Concentrated fatty acid sources offer the advantage of a reduced need for preconcentration of the feedstock in chemical processes. Safflower (*Carthamus tinctorius*) oil contains about 78 percent linoleic acid, and some high-oleic-acid type sunflower seed oils have been claimed to contain over 92 percent oleic acid.

Short-chain fatty acid sources mainly have been limited to coconut and palm kernel oils which contain C8, C10, and C12 fatty acids, but only at levels of 6 to 10 percent. *Cuphea* varieties, containing about 33 percent oil, are being evaluated as sources of short chain fatty acids, with reports of 18 percent C10 and 57 percent C12 in the oil of *C. carthagenesis*, 73 percent C8 and 24 percent C10 in *C. painteri* oil, 3 percent C8 and 87 percent C10 in *C. ignea* oil, and 85 percent C10 in *C. llavea*.[38]

Long-chain fatty acids generally are undesirable for food and feed use, but have been in demand for production of lubricants and polymers (Nylon 1313), and as plasticizers. Potential sources include: *Crambe abyssinica* oil containing over 60 percent C22; *Limnanthes alba* (Meadowfoam) oil with 95 percent C20 and C22; *Lunaria annua* oil wth 40 percent C22 and 20 percent C24; *Selenia grandis* oil with 58 percent C22; *Leavenworthia alabamica* oil with 50 percent C22; and *Marshallia caespitosa* oil with 44 percent C22.[38]

Hydroxy fatty acids are used for thickening greases and for improving the pliability of plastic covering materials. Domestic needs have been met primarily by imported castor oil. *Lesquerella* species are being explored for this purpose, with 70 percent 14-OH-C20 reported for *L. gracilis* oil, and 50 percent 12-OH-C18 diene for *L. densipila* oil.[38]

Epoxy fatty acids are used in the production of plastics and coatings. Significant quantities of soy oil are epoxidized yearly using a relatively expensive process. Potential new sources include: *Vernonia pauciflora* and *V. anthelmintica* oils, with 73 to 80 percent and 68 to 75 percent epoxy acid contents, respectively; *Euphorbia lagascae* with 60 to 70 percent epoxy acid content; and Stokes aster *Stokesia laevis* oil with 75 percent epoxy acid content.[38]

Conjugated Unsaturation Sources. Multiple bonds that are separated from each other by one single bond are said to be conjugated.

$$CH_3(CH_2)_n{-}CH{=}CH{-}CH{=}CH{-}(CH_2)_nCOOH$$

Conjugated double bonds

Fatty acids with conjugated unsaturation frequently are desired as chemical intermediates for industrial products. Tung oil is a source of conjugated unsaturation, which also is obtained by alkali isomerization of soybean and linseed fatty acids. Plant oils that have been reviewed as potential new sources of conjugated unsaturation include: *Valeriana officinalis* with 40 percent of its oil in 9, 11, 13 unsaturation; *Calendula oficinalis* oil with 55 percent 8, 10, 12 unsaturation; *Centranthus macrosiphon* oil with 65 percent 9, 11, 13 unsaturation; and *Impatiens edgeworthii* oil with 60 percent 9, 11, 13, 15 unsaturation.

Drying Oils. The value of drying oils depends on their ability to polymerize or dry upon application onto a surface to form tough, adherent, impervious, and abrasion-resistant films. These film-forming properties are closely related to the degree of unsaturation, and linsed, castor, fish, soybean, tung, oiticia, and tall oils are some of the important sources of domestic drying oils.

FATS AND OILS DEGRADATION AND ANTIOXIDANTS

The major types of fats and oils degradation are *lipolysis* (liberation of free fatty acids) and *oxidation* (formation of peroxides and aldehydes). Lipolysis can occur as the result of chemical (acid, base or water) or enzymatic (lipase) action. For example, free fatty acids develop more rapidly in oilseeds stored at increased moisture contents and temperatures because of increased enzymatic activity. Also, lipolysis in coconut oil can be induced by moisture and heat alone. Lipolytic spoilage and oxidative spoilage sometimes are called *hydrolytic rancidity* and *oxidative rancidity*, respectively. In some industries (such as dairy products) "rancidity" almost exclusively means lipolytic spoilage, whereas in the fats and oils industry "rancidity" often means oxidative spoilage.

Oxidation is thought to be induced by (singlet) oxygen at unsaturated bonds of fatty acids, and may be accelerated by the type of fatty acid, light, heat, oxygen concentration, extended surface areas exposed to air, and the presence of prooxidants (cobalt, copper, iron, manganese, and nickel), and by specific water activity (A_w) levels of 0.1 to 0.3 and 0.55 to 0.85. Fatty acids in triacylglycerols can undergo oxidation, but the reaction is more rapid if they have been freed by lipolysis. The process of oxidation consists of free radical creation, which then leads to the formation of hydroperoxides whose decomposition products catalyze a self-accelerating autoxidation process.

$$\begin{aligned}
&RH + \text{singlet oxygen} \rightarrow ROOH\\
&ROOH + \text{catalysts} \rightarrow RO^{\bullet} + OH^{-}\\
&RO^{\bullet} + RH \rightarrow ROH + R^{\bullet}\\
&\\
&R^{\bullet} + O_2 \rightarrow ROO^{\bullet}\\
&ROO^{\bullet} + RH \rightarrow ROOH + R^{\bullet}\\
&R^{\bullet} + O_2 + RH \rightarrow ROO^{\bullet} + R^{\bullet}\\
&\\
&R^{\bullet} + R^{\bullet} \rightarrow \text{Nonradical product}\\
&ROO^{\bullet} + R^{\bullet} \rightarrow \text{Nonradical product}\\
&ROO^{\bullet} + ROO^{\bullet} \rightarrow \text{Nonradical product}
\end{aligned}$$

The extent of prior oxidative activity in a fat sample may be estimated by the following: peroxide value (PV), thiobarbituric acid test (TBA) (which measures the presence of malonaldehyde), anisidine value—Totox, Kreis test, oxirane test, total and volatile carbonyl

compounds, chromatographic analysis, ultraviolet spectroscopy, fluorescence, and especially organoleptic evaluation. However, the reader should be cautioned that the resulting peroxides are unstable and decompose. Thus, peroxide values are not cumulative and are not always indicative of the extent of earlier oxidation.

Predisposition (stability) to oxidation can be estimated for fats by the active oxygen method (AOM), and for fat-containing products by the Schall oven test and the oxygen bomb method and as a function of oxygen absorption.

Fats and fat-containing products may be stabilized against oxidation by the addition of antioxidants—compounds believed to inhibit an autoxidation chain reaction by acting as hydrogen donors or free radical acceptors. The major food grade synthetic antioxidants used include butylated hydroxytoluene (BHT), tertiary butylhydroquinine (TBHQ), propyl gallate (PG), and 2,4,5-trihydroxybutyrophenone (THBP), and generally they are allowed in food products at 0.01 percent of total fat weight each, with a total of 0.02 percent if a combination is used. Synergists used in combination with antioxidants include citric acid, isopropyl acid, phosphoric acid, ascorbic acid, ascorbyl palmitate, iso-ascorbic acid, tartaric acid, and lecithin, most of which act as chelators of prooxidative metals. Natural antioxidants include gum guaiac, tocotrienols (including vitamin E), and oil of rosemary (containing rosmaridiphenol). Antioxidants that no longer are used in food systems, or have not come into use, include the synthetics butylatedhydroxy anisole (BHA), 4-hydroxymethyl-2,6-ditertiarybutylphenol (an analog of BHT), and the natural compound nordihydroquaiaretic acid (NDGA) extracted from the desert plant *Larrea divaricata.*

The crosslinking that occurs in drying oils also is an oxidation reaction. Polymerization of oils can occur rapidly in products such as fish meal that contain significant amounts of polyunsaturated triacylglycerols—even to the point of building up sufficient heat for a pile to spontaneously burst into flame. Thus, feed grade antioxidants such as ethoxyquin often are added before storage. Ethoxyquin also has been used for stabilizing feedstuffs against oxidation, thus sparing the natural tocopherols for their vitamin E activity in nutrition.

Butylated hydroxytoluene (BHT)

Tertiary butylhydroquinone (RBHQ)

Propyl gallate (PG)

Butylated hydroxyanisol (BHA)

Trihydroxybutyrophenone (THBP)

Hydroxymethylditertiarybutylphenol

Synthetic and natural antioxidants are known to be heat- and steam-distillable, and preferably should be added to oil products after the maximum heat encountered in processing. It further is known that many natural phenolic compounds in cereal-oilseed products as well as reaction products of Maillard (nonenzymatic amino acid–reducing sugar) browning and natural wood smoke have antioxidant effects. Additional techniques for reducing oxidative activity include maintaining nitrogen blankets on oils stored in tanks, and preferentially distributing oils in opaque containers or brown bottles.

ANALYTICAL METHODS

Selected analytical methods, adopted by the American Oil Chemists' Society (AOCS) for characterizing composition, structure, physical properties, and stability of fats and oils, are summarized below.[39] Prescribed equipment must be used and conditions followed. Some of these techniques are limited to specific oil species, but adaptations are available for other species. Procedures for sample drawing and preparation also are specified.

Active Oxygen Method for Fat Stability—AOM (Cd 12–57): determines the time (in hours) for a sample of fat or oil to attain a predetermined peroxide value under conditions of the test. The method is used to estimate the comparative oxidative stability of fats and oils.

Bleaching Test for Soybean Oil (Cc 8e–63): determines the color of a sample of soybean oil after treatment with a specified bleaching earth. Specific methods exist for other oil species.

Cloud Point Test (Cc 6–25): determines the temperature at which a cloud first forms in a sample of melted fat in the first stage of crystallization.

Cold Test (Cc 11–53): measures the relative resistance of a sample to crystallize at an established temperature in terms of time. The test is often used as an index of stearin removal in winterization of salad oils.

Color Measurement by Lovibond—Wesson (Cc 13b–45): determines the color of clear oil samples by comparison with glasses of known color characteristics.

Fatty Acid Composition by Gas Chromatography (Ce 1–62): quantitatively determines saturated and unsaturated fatty acids with 8 to 24 carbon atoms in animal fats, vegetable oils, marine oils, and fatty acids after conversion to their methyl ester forms.

Fatty Acid Composition by GLC (Ce 1c–89): measures fatty acid composition and levels of *trans* unsaturation and *cis*, *cis* methylene-interrupted unsaturation of vegetable oils using capillary gas liquid chromatography.

Flash Point—Pensky-Martens Closed Cup for Fats and Oils (Cc 9b–55): determines the temperature at which a sample will flash when a test flame is applied. This technique is sometimes used to estimate levels of residual hexane as well as to assure safety of workers handling the oil.

Free Fatty Acids in Crude and Refined Fats and Oils (Ca 5a–40): determines FFAs, as oleic acid, by ethanolic sodium hydroxide titration.

Halphen Test for Detecting Cottonseed Oil (Cb 1–25): estimates the presence of cottonseed oil in vegetable or animal fats or oils as the result of a pink color formed between the reagent and cyclopropenoic fatty acids (sterculic and malvalic) normally present in cottonseed oil.

Hexane Residues in Fats and Oils (Ca 3b–87): determines, by gas chromatography, the "free" volatile hydrocarbons remaining in fats and oils after extraction with hydrocarbon solvents. The results are expressed in terms of hexane.

Insoluble Impurities in Fats and Oils (Ca 3a–46): determines dirt, meal, and foreign substances that are insoluble in kerosene and petroleum ether.

Iodine Value of Fats and Oils—Wijs (Cd 1–25): measures the unsaturation of fats and oils in terms of centigrams of iodine absorbed per gram of sample. The method is applicable to all normal fats that do not contain conjugated double bonds. It often is used to estimate the degree of hydrogenation of oils.

Lard, Foreign Fat—Boemer Number (Cb 5–40): estimates the presence of tallows and similar fats, based on differences in melting points of foreign glycerides and fatty acids as compared to pure pork fat.

Melting Point—Capillary Tube Method (Cc 1–25): determines the temperature at which a sample of fat in a closed capillary becomes completely clear and liquid.

Melting Point—Mettler Dropping Point (Cc 18–80): determines the temperature at which a sample becomes sufficiently fluid to flow in a specified apparatus.

Melting Point, Open Tube—Softening Point (Cc 3–25): determines the temperature at which a solidified fat, in an open capillary tube, softens sufficiently to slip and rise to the top of the heating bath. This method is applicable to fats such as coconut oil, stearin, hydrogenated fats, and hard tallows. The results sometimes are reported as "melting slip point," but the method is different from the AOCS Slip Point (Cc 4–25).

Melting Point—Wiley Method (Cc 2–38): determines the temperature at which a sample disc of solidified fat assumes a spherical shape while suspended in a heating bath with an alcohol–water density gradient.

Moisture—Distillation Method (Ca 2a–45): determines only moisture in triacylglycerols and emulsions by distillation with an immiscible solvent (toluene).

Moisture and Volatile Matter—Air Oven Method (Ca 2c–25): determines moisture and volatile matter by heating in a hot air oven. This method is applicable to animal and vegetable fats, but not to drying oils, coconut group fats, or oils with added monoacylglycerols.

Oil (Aa 4–38): determines oil content in a dried sample of oil-bearing material by extraction with petroleum ether. This method is specific for cottonseed, which first must be fumed with hydrochloric acid to prevent oil adsorption to the fiber. Additional methods exist for other oilseeds.

Peroxide Value, Fats and Oils (Cd 8–53): determines all substances, in terms of milliequivalent of peroxide per 1000 grams of sample, that oxidize potassium iodide (KI). These substances generally are assumed to be peroxides or products of fat oxidation.

Phosphorus in Oils (Ca 13–55): estimates the phospholipid content of crude, degummed, and refined vegetable oils in terms of phosphorus.

Refining Loss, Vegetable Oils, Crude (Ca 9a–52): determines the loss of free fatty acids and impurities when crude oils are refined under specified procedures.

Residual Lint (Aa 7–55): determines the lint content of cottonseed by fuming (digesting) with hydrochloric acid.

Saponification Value of Fats and Oils (Cd 3–25): determines the number of milligrams of (alcoholic) potassium hydroxide necessary to saponify a one-gram sample of a fat or an oil.

Smoke, Flash, and Fire Point—Cleveland Open Cup Method (Cc 9a–48): determines the temperatures at which fats and oils smoke, flash, or burn. Smoke point determinations sometimes are used to follow degradation of frying oils with use.

Solid Fat Content (SFC) of Fats and Oils by NMR (Cd 16–81): estimates the percent of solids in a semisolid fat on the basis of the pulsed nuclear magnetic resonance signal of hydrogen in the liquid fraction.

Solid Fat Index (SFI)—Dilatometric Method (Cd 10–57): estimates the percent of solids in a semisolid fat on the basis of

changes in volume with temperature. This method utilizes glass dilatometers.

Triglycerides by GLC (Ce 5b–86): quantitatively determines triglycerides (triacylglycerols) in liquid vegetable oils in terms of molecular weight and degree of unsaturation as a function of their equivalent carbon number using high-pressure liquid chromatography.

Unsaponifiable Matter in Fats and Oils, Including Marine Oils (Ca 6b–53): determines substances dissolved in fats and oils that cannot be saponified (turned into sodium salts) by usual caustic treatment, including higher aliphatic alcohols, sterols, pigments, and hydrocarbons. This method is not suitable for marine oils or feed grade fats.

The following characterization methods also are used, but have not been included in AOCS *Official Methods*:

Anisidine Value: measures the extent of oxidative deterioration, which spectrophotometrically estimates aldehydes, and secondary oxidation products.

Schaal Test (Schaal Oven Method): an accelerated test for determining oxidative stability of a fat or a fat-containing food product. Results are reported as the time elapsed until a rancid odor is detected in a covered glass container held at an elevated temperature.

Totox Value: an estimate of the degree of oxidation of a fat or oil, calculated as: Totox = 2(Peroxide Value) + Anisidine Value.

TRADING OF FATS AND OILS

The Federal Grain Inspection Services certifies the quality of whole oilseeds for export trading. Fats and oils are traded domestically and in export markets, by mutual agreement of both parties, under trading rules established by various groups, including: National Oilseed Processors Association, Washington, DC—soybean oil, meal, and lecithin; National Cottonseed Products Association, Memphis, TN—cottonseed, oil, soapstock, tank bottoms, cottonseed cake and meal, cottonseed hulls, linters, and peanut oil; National Institute of Oilseed Products, Monterey, CA—coconut oil, palm oil, palm kernel oil, oleic safflower oil, high oleic sunflower oil, corn oil, LEAR rapeseed oil, raw soybean oil, cottonseed oil, peanut oil, sesame oil, teaseed oil, tung oil, *plus* palm kernels, babassu seed, flaxseed, castor seeds, ground nuts (peanuts), sesame seeds, safflower seed; and American Fats and Oils Association, Inc., Greatneck, NY—animal tallow and grease, vegetable oils, animal and poultry proteins, and fishmeal.

REFERENCES

1. Ege, S. N., *Organic Chemistry*, D. C. Heath and Company, Lexington, MA, 1984.
2. National Academy of Sciences, *Recommended Dietary Allowances*, 10th ed., Washington, DC, 1989.
3. Pryde, E. H., *Fatty Acids*, American Oil Chemists' Society, Champaign, IL, 1979.
4. Pryde, E. H., Princen, L. H., and Mukherjee, K. D. *New Sources of Fats and Oils*, American Oil Chemists' Society, Champaign, IL, 1981.
5. Cahn, R. S., Ingold, C. K., and Prelog, V., "The Specification of Asymmetric Configuration in Organic Chemistry," *Experientia*, **12**, 81–94 (1956).
6. Hirschman, H., "The Nature of Substrate Asymmetry in Stereoselective Reactions," *J. Biol. Chem.*, **235**, 2762–2767 (1960).
7. Nawar, W. W., "Lipids," in *Food Chemistry*, 2nd ed., O. R. Fennema (Ed.), pp. 139–244, Marcel Dekker, Inc., New York, 1985.
8. Hagemann, J. W., Tallent, W. H., Barve, J. S., Ismail, I. A., and Gunstone, F. D., "Polymorphism in Single-Acid Triglycerides of Positional and Geometric Isomers of Octadecanoic Acid," *J. Am. Oil Chem. Soc.*, **52**, 204–207 (1975).

9. Rossell, J. B., "Phase Diagrams of Triglyceride Systems," in *Advances in Lipid Research*, Vol. 5, R. Paoletti and D. Kritchevsky (Eds.), pp. 353–408. Academic Press, New York, 1967.
10. Small, D. M., *The Physical Chemistry of Lipids: Handbook of Lipid Research, 4*, Plenum Press, New York, 1986.
11. Hernqvst, L., "Chocolate Temper," in *Industrial Chocolate Manufacture and Use*, S. T. Beckett (Ed.), pp. 153–171, Avi-Van Nostrand Reinhold, New York, 1988.
12. Hernqvst, L., "Crystal Structures of Fats and Fatty Acids," in *Crystallization and Polymorphism of Fats and Fatty Acids*, N. Garti and K. Sato (Eds.), pp. 97–137, Marcel Dekker, Inc., New York, 1988.
13. Shipley, C. G., "X-Ray Crystallographic Studies of Aliphatic Lipids," in *Handbook of Lipid Research 4: The Physical Chemistry of Lipids*, D. M. Small (Ed.), pp. 97–147, Plenum Press, New York, 1986.
14. Hagemann, J. W., "Interesterification Products and Processes," *J. Am. Oil Chem. Soc.*, **44**, 414A–456A (1969).
15. Poot, C., Dijkshoorn, W., Haigton, A. J., and Verburg, C. C., "Laboratory Separation of Crystals from Plastic Fats Using Detergent Solution," *J. Am. Oil Chem. Soc.*, **52**, 69–72 (1975).
16. Rivarola, G., Anon, M. C., and Calvelo, A., "Crystallization of Waxes during Sunflowerseed Oil Refining," *J. Am. Oil Chem. Soc.*, **62**, 1508–1513 (1986).
17. Rivarola, G., Anon, M. C., and Calvelo, A., "Crystallization of Waxes in Sunflowerseed Oil," *J. Am. Oil Chem. Soc.*, **65**, 1771–1513 (1988).
18. Technical Committee—ISEO, *Food Fats and Oils*, Institute of Shortening and Edible Oils, Inc., Washington, DC, 1988.
19. Lusas, E. W., Watkins, L. R., and Rhee, K. C., "Separation of Fats and Oils by Solvent Extraction: Non-traditional Methods," in *World Conference Proceedings, Edible Fats and Oils Processing: Basic Principles and Modern Practices*, D. R. Erickson (Ed.), pp. 56–79, American Oil Chemists' Society, Champaign, IL, 1990.
20. Cavanagh, G. C., "Neutralization II. Theory and Practice of Non-conventional Caustic Refining by Miscella Refining and by the Zenith Process," in *World Conference Proceedings, Edible Fats and Oils Processing: Basic Principles and Modern Practices*, D. R. Erickson (Ed.), pp. 101–106, American Oil Chemists' Society, Champaign, IL, 1990.
21. Martinez Suarez, J. M., and Mendoza, J. A., "Olive Oil Processing and Related Aspects. Extraction and Refining: General Aspect," in *Proceedings World Conference on Emerging Technologies in the Fats and Oils Industry*, A. R. Baldwin (Ed.), pp. 299–305, American Oil Chemists' Society, Champaign, IL, 1968.
22. Hastert, R. C., "Cost/Quality/Health: The Three Pillars of Hydrogenation," in *World Conference Proceedings, Edible Fats and Oils Processing: Basic Principles and Modern Practices*, D. R. Erickson (Ed.), pp. 142–151, American Oil Chemists' Society, Champaign, IL, 1990.
23. Hvolby, A., "Expansion of Solidifying Fats," *J. Am. Oil Chem. Soc.*, **51**, 50–54 (1974).
24. Segers, J., and van de Sande, R., "Degumming—Theory and Practice," in *World Conference Proceedings—Edible Fats and Oils Processing: Basic Principles and Modern Practices*, D. R. Erickson (Ed.), pp. 88–93, American Oil Chemists' Society, Champaign, IL, 1990.
25. Hendrix, B., "Neutralization I.—Theory and Practice of Conventional Caustic (NaOH) Refining," in *World Conference Proceedings—Edible Fats and Oils Processing: Basic Principles and Modern Practices*, D. R. Erickson (Ed.), pp. 94–101, American Oil Chemists' Society, Champaign, IL, 1990.
26. Mag, T. K., "Bleaching—Theory and Practice," in *World Conference Proceedings, Edible Fats and Oils Processing: Basic Princples and Modern Practices*, D. R. Erickson (Ed.), pp. 107–116, American Oil Chemists' Society, Champaign, IL, 1990.
27. Loft, S. C., "Deodorization—Theory and Practice," in *World Conference Proceedings, Edible Fats and Oils Processing: Basic Principles and Modern Practices*, Erickson, D. R. (Ed.), pp. 117–123, American Oil Chemists' Society, Champaign, IL, 1990.
28. Young, F. V. K., "Physical Refining," in *World Conference Proceedings, Edible Fats and Oils Processing: Basic Principles and Modern Practices*, D. R. Erickson (Ed.), pp. 124–136, American Oil Chemists' Society, Champaign, IL, 1990.
29. Kritchevsky, D., "*Trans* Unsaturated Fat in Nutrition and Health," in *World Conference Proceedings—Edible Fats and Oils Processing: Basic Principles and Modern Practices*, D. R. Erickson (Ed.), pp. 158–165, American Oil Chemists' Society, Champaign, IL, 1990.
30. Rozendaal, A., "Interesterification of Oils and Fats," in *World Conference Proceedings—Edible Fats and Oils Processing: Basic Principles and Modern Practices*, D. R. Erickson (Ed.), pp. 152–157, American Oil Chemists' Society, Champaign, IL, 1990.
31. Going, L. H., "Interesterification Products and Processes," *J. Am. Oil Chem. Soc.*, **44**, 414A–456A (1969).
32. Tirtiaux, A., "Dry Fractionation: A Technique and an Art," in *World Conference Proceedings, Edible Fats and Oils Processing: Basic Principles and Modern Practices*, D. R. Erickson (Ed.), pp. 136–141, American Oil Chemists' Society, Champaign, IL, 1990.
33. Shahidi, F., "Rapeseed and Canola: Global Production and Distribution," in *Canola and Rapeseed: Production, Chemistry, Nutrition and Processing Technology*, Avi-Van Nostrand Reinhold, New York, 1990.

34. Becher, P., and Schick, M. J., "Macroemulsions," in *Nonionic Surfactants: Physical Chemistry*, M. J. Schick (Ed.), pp. 435–491, Marcel Dekker, Inc., New York, 1987.
35. Horrocks, L. A., "Nomenclature and Structure of Phosphatides," in *Lecithins: Sources, Manufacture and Uses*, B. F. Szuhaj (Ed.), pp. 1–6, American Oil Chemists' Society, Champaign, IL, 1989.
36. Dziezak, J. D., "Fats, Oils and Fat Substitutes," *Food Technol.*, **43**(7), 65–74 (1989).
37. Staff Article, "Fat Substitute Update," *Food Technol.*, **44** (3), 92, 94, 97 (1990).
38. Princen, L. H., *Crop Resources*, Academic Press, Inc., New York, 1977.
39. AOCS, *Official Methods and Recommended Practices of the American Oil Chemists' Society*, 4th ed., American Oil Chemists' Society, Champaign, IL, 1989.

SUPPLEMENTAL READING

American Fats and Oils Association, Inc., *Trading and Arbitration Rules*, Great Neck, NY, 1990.

Baldwin, A. R. (Ed.), *Proceedings of the World Conference on Emerging Technologies in the Fats and Oil Industry*, American Oil Chemists' Society, Champaign, IL, 1986.

Baldwin, A. R. (Ed.), *Proceedings of the Seventh International Conference on Jojoba and Its Uses*, American Oil Chemists' Society, Champaign, IL, 1989.

Beare-Rogers, J. (Ed.), *Dietary Fat Requirements in Health and Development*, American Oil Chemists' Society, Champaign, IL, 1988.

Bernardidi, E., *Oilseeds, Oils and Fats*, Vol. I, *Raw Materials*, and Vol. II, *Oils and Fats Processing*, B. E. Oil, Rome, Italy, 1983.

Emken, E. A., and Dutton, H. J. (Eds.), *Geometrical and Positional Fatty Acid Isomers*, American Oil Chemists' Society, Champaign, IL, 1986.

Erickson, D. R. (Ed.), *World Conference Proceedings Edible Fats and Oils Processing: Basic Principles and Modern Practices*, American Oil Chemists' Society, Champaign, IL, 1990.

Erickson, D. R., Pryde, E. H., Brekke, O. L., Mounts, T. L., and Falb, R. A. (Eds.), *Handbook of Soy Oil Processing and Utilization*, American Oil Chemists' Society, Champaign, IL, 1980.

Garti, N., and Sato, K. (Eds.), *Crystallization and Polymorphism of Fats and Fatty Acids*, Marcel Dekker, Inc., New York, 1988.

Hastert, R. A. (Ed.), *Hydrogenation: Proceedings of an AOCS Colloquium*, American Oil Chemists' Society, Champaign, IL, 1987.

International Trade Publications Ltd., *Oils and Fats International Directory Survey*, England, 1990.

Kabara, J. J. (Ed.), *The Pharmacological Effect of Lipids III*, American Oil Chemists' Society, Champaign, IL, 1990.

Larrson, K. and Frieberg, S. W. (Eds.), *Food Emulsions*, Marcel Dekker, Inc., New York, 1990.

Min, D. B., and Smouse, T H. (Eds.), *Flavor Chemistry of Lipid Foods*, American Oil Chemists' Society, Champaign, IL, 1989.

National Cottonseed Products Association, *Trading Rules*, Memphis, TN, 1990–91.

National Institute of Oilseed Products, *Trading Rules*, Monterey, CA, 1990.

National Oilseed Processors Associations, *Yearbook & Trading Rules*, Washington, DC, 1990–91.

"Proceedings of World Conference on Processing of Palm, Palm Kernel, and Coconut Oils," *J. Am. Oil Chem. Soc.*, **62**(2), 139–468 (1985).

Soyatech, Inc., *Soya Bluebook*, Bar Harbor, ME, 1990.

Pryde, E. H. (Ed.), *Fatty Acids*, American Oil Chemists' Society, Champaign, IL, 1979.

Pryde, E. H., Princen, L. H., and Mukherjee, K. D. (Eds.), *New Sources of Fats and Oils*, American Oil Chemists' Society, Champaign, IL, 1981.

Sevanian, A. (Ed.), *Lipid Peroxidation in Biological Systems*, American Oil Chemists' Society, Champaign, IL, 1988.

Shahidi, F. (Ed.), *Canola and Rapeseed: Production, Chemistry, Nutrition, and Processing Technology*, Avi-Van Nostrand Reinhold, New York, 1990.

Shick, M. J. (Ed.), *Nonionic Surfactants: Physical Chemistry*, Marcel Dekker, Inc., New York, 1987.

Small, D. M. (Ed.), *Handbook of Lipid Research 4: The Physical Chemistry of Lipids*, Plenum Press, New York, 1986.

Stansby, M. E. (Ed.), *Fish Oils in Nutrition*, Avi-Van Nostrand Reinhold, New York, 1990.

Swern, D. (Ed.), *Bailey's Industrial Oil and Fat Products*, John Wiley & Sons, New York, 1979.

Szuhaj, B. F. (Ed.), *Lecithin: Sources, Manufacture, Uses*, American Oil Chemists' Society, Champaign, IL, 1989.

Wisniak, J. (Ed.), *The Chemistry and Technology of Jojoba Oil*, American Oil Chemists' Society, Champaign, IL, 1987.

9

Sugar and Other Sweeteners

Raymond D. Moroz* and Charles D. Broeg**

INTRODUCTION

Sugar and starch are among those organic "chemicals" found so abundantly in nature that no serious efforts have been made to synthesize them commercially from coal (or petroleum), air, and water. Both are available at such concentrations in some plants that sizable industries have resulted from growing those plants and extracting carbohydrates therefrom.

The primary use of sugar is in the manufacture of food or as a food in itself. When used for such purposes, most of it is highly refined or purified, but considerable quantities are consumed in some areas as a crude product, as in India where about 10 percent of the matured cane is consumed as unmilled sugar (gur and khandsari). Sugar is used to a limited extent in the production of other chemicals such as sucrose esters and in the form of by-product molasses as a substrate for fermentation processes. Currently, the juice of sugar-producing plants also is utilized for fermentation.[1]

Starch is extensively used in the textile and paper industries, but food is also a major outlet, especially if one considers flours that are primarily starch as a crude form of starch. However, a very important outlet for starch is its conversion to dextrose and glucose syrups for use in the food industry. Technological developments and governmental price-support programs for the domestic (U.S.) cane and beet sugar industries have greatly magnified the importance of starch as a source of sugars.[2] Those developments are based on the enzymatic conversion of glucose to fructose. The following discussion explores briefly the sources of various sugars and the processes by which they are converted into commercial products.

SUGAR

History

The ancestry of the sugar cane and its use as a food have been traced to the island of New Guinea. (However, a very interesting legend related by J. A. C. Hugill in Chapter 1 of *Sugar: Sucrie and Technology*[3] associates sugar cane with the origin of the human race.)

*Crompton & Knowles Corporation.
**Sugar Consultant.

Around 8000 B.C., the plant started on its migration from New Guinea to the many areas of southeastern Asia, Indonesia, the Philippines, Malay, Indochina, and eastern India, with humans probably as its main dispersal agent.[4] It was in Bengalese India that sugar cane first was cultivated as a field crop and that the juice was manufactured into various solid forms. A general knowledge of sugar was prevalent throughout India by 400 B.C. By the tenth century A.D., sugar cultivation and manufacturing had become important industries in Persia and Egypt. The early Islamic movement spread the knowledge of the sugar industry throughout the Mediterranean area. On the second voyage of Columbus to America, in 1493, sugar cane was introduced in Santo Domingo. It spread rapidly through the West Indies and Central America. Cortez brought the cane to Mexico, and Pizarro introduced it in Peru. By 1600, the sugar industry was the largest in tropical America.[5]

The modern sugar industry dates from the end of the eighteenth century when steam replaced animal energy and made possible the development of larger and more efficient production units. The vacuum pan appeared in 1813, bag filters in 1824, multiple-effect evaporators in 1846, filter presses in 1850, centrifugals in 1867, dryers in 1878, and packaging machines in 1891.[6]

Cultivation of the sugar beet plant and the manufacture of sugar from the beet developed in the industrial nations of Europe within the last two centuries. In 1747, a German chemist, Marggraf, established that sugar from beets was the same as sugar from cane. His pupil, Achard, in 1799, demonstrated that sugar can be commercially prepared from beets. During the Napoleonic wars, a short-lived beet sugar industry was established in France. It was slowly revived 20 years after Waterloo and spread to most European countries.[7] Today, sugar from beets accounts for about 35 percent of the world's supply.

Cane Sugar

Figure 9.1 is a generalized flow diagram for the manufacture of raw cane sugar.[8] The several steps are described in the sections that follow.

Agriculture. The sugar cane is a large perennial tropical grass belonging to the genus *Saccharum.* There are three basic species, *S. officinarum*, *S. robustum*, and *S. spontaneum*, and a large number of varieties. Sugar cane is propagated commercially by cuttings, each cutting consisting of portions of the cane plant having two or more buds. The buds sprout into shoots from which several other shoots arise below the soil level to form a clump of stalks or a "stool." From 12 to 20 months are required for crops maturity from new plantings and about 12 months for ratoon crops (i.e., cane stalks arising from stools that have been previously harvested). Most of the field operations have been mechanized, but planting is still done by hand in some cane-producing areas. Fields are replanted after two to five "cuttings" have been made from the original plantings. As cane grows, its foliage is largely limited to the upper one-third of the plant (the lower leaves die off from shading as the stalk grows taller). The cane stalk is round and from a few to more than 10 feet long when mature. It is covered with a hard rind that is light brown

Fig. 9.1. Flow diagram showing principal steps in sugar cane manufacture. (*Adapted from Clarke, Margaret E., "Cane Sugar Process and Byproduct Streams," Iowa Biotechnology Consortium, Mar. 1991.*)

or green, yellowish green, or purple in color, depending on the variety. The stalk consists of a series of joints or internodes separated by nodes. The rind and the nodes are of a woody nature, whereas the internode is soft pith. The internode contains the greater part of the juice containing sucrose and the nodes contain "eyes" or buds that "sprout" when planted.[9]

Harvesting. The sugar cane is still cut by hand with machete-type knives in many producing areas. The canes are cut at ground level, and, at the same time, leaves and tops of the stalk are removed. In areas where labor is scarce or expensive, machine harvesting has come into widespread use. Some harvesters cut, chop, and load the cane into transporting vehicles in one operation at rates of 30 to 45 tons/hr. Other harvesters cut, bundle, and dump cane in rows for pickup by other machines. Transportation of field cane to the mill is accomplished by railcars, trucks, trailers, and carts.

Preparation of Cane for Milling. Mechanically harvested cane must be washed before milling in order to eliminate soil, rocks, and field trash. In some instances, "dry cleaning" precedes washing. Washing systems range from a simple spraying with warm water on the carrier or table to an elaborate system consisting of conveyors with water jets, stripping rolls, and baths for removal of stones.

Juice Extraction. The cane is first prepared for grinding by one or more of the following operations: (1) chopping it into smaller pieces with one or two sets of rotating knives (400–600 rpm; see Fig. 9.2); or (2) disintegrating the cane into finer sizes by a crusher and/or shredder (Fig. 9.3). The crusher usually is a two-roller mill of the Krajewski or Fulton type.

Recently, a new machine has been developed that combines the action of the chopper and the shredder into a single unit. It is commonly called a Unigrator or the "Ducasse Shredder."[10,11] The advantages gained have been a greater increase in juice extraction by the mills, an increase in the grinding rate, and, in some cases, a decrease in power consumption by the milling operation.

Radically new approaches to the preparation of sugar cane for the extraction of sugar have been under development for several years.[12] The basic concept consists of chopping cane stalks into short segments (8–12 in. long) and feeding these segments into a machine (Figs. 9.4 and 9.5) that separates the cane into three major fractions: (1) epidermis, (2) rind, and (3) pith. This approach to the preparation of cane for the extraction of sugar have several potential advantages: (1) reduced power requirements, (2) simplification of juice ex-

Fig. 9.2. Farrel Type K-4 cane knife to handle 4500 tons of cane per 24-hour day. (*Courtesy Farrel Co., Division of USM Corp.*)

Fig. 9.3. Fulton crusher roll. (*Courtesy Fulton Iron Works Co., St. Louis, MO.*)

traction processes, and (3) the recovery of not less than three by-products having greater utility than bagasse.[12]

The crushed and shredded cane then passes through a series of three horizontal rollers (mill) arranged in a triangular pattern with the top roll rotating counterclockwise (Fig. 9.6) and the bottom two rollers clockwise. A series of three-roller mills, numbering 3 to 7, is called a tandem (Fig. 9.7). The pressure on the top roll is regulated by hydraulic rams and averages about 300 tons. Below each mill there is a juice pan into which expressed juice flows. The crusher and first mill extract 60 to 70 percent of the cane juice, and the remaining mills take out 22 to 24 percent.

When the fiber content of the bagasse reaches about 50 percent, extraction by conventional milling approaches zero. Because the juice that remains with the fiber contains the same proportion of sucrose as the original juice of the cane, the yield of sugar would be substantially lowered if extraction were terminated at that point. Consequently, a process called "compound imbibition" is used to reduce the sucrose in the fiber by repeated dilution and milling. In a five-mill tandem, water is added to the fourth mill, and the expressed juice from that mill is brought back to the second mill. The expressed juice from the third mill is recirculated to the first mill and the fifth mill to the third mill. In this way, the juice in the bagasse is always diluted ahead of each mill. The amount of imbibition applied at each mill is approximately equal to the amount of water applied to the fourth mill or to the penultimate mill in a differently numbered tandem.[13]

Before the expressed juice goes to clarification, it is screened through perforated metal screens with openings about 1 mm in diameter. Additional screening may be provided by stationary or vibrating metal cloths.

Diffusion. Diffusion is used to extract juice from the cells in the sugar cane. The diffusers used for this have capacities that range from 1000 to 6000 tons of cane per day. Normally they are used in conjunction with part of a milling tandem, the crusher and the first mill. Preparation of the cane prior to its entry into the diffuser is essential for good extraction. In the diffuser, the crushed cane is countercurrently washed with imbibition water at temperatures ranging from 50°C to 75°C. The last stage of a diffuser, the bagasse exit, receives water that gradually increases in sugar content as it proceeds to the first stage. Most of the sugar is extracted in the first four or five stages by simple displacement of sucrose from the open ruptured cane cells. In later stages, diffusion appears to take place in unruptured cane cells. One or two mills are used to express water from the bagasse after discharge from the diffuser. The crusher and one mill extract 70 percent of the juice from the cane, and the diffuser removes 30 percent of the expected yield. Sucrose extraction from a mill-diffusion system averages 97 percent compared to 90 to 91 percent with a straight milling system.[14]

Since 1962, a number of diffusers have been installed in cane mills throughout the world. Cane diffusers now are manufactured by a number of companies including Braunschweigische Maschinenbauanstalt, CF&I Engineers Inc., De Danske Sukker-

Fig. 9.4. Tilby separator process. (*Courtesy Ander-Cane, Inc., Naples, FL.*)

fabrikker, De Smet, S. A., Sucatlan Engineering, and Suchem Inc. of Puerto Rico (Figs. 9.8, 9.9).

Juice Purification. The juice from the milling station is an acidic, opaque, greenish liquid containing soluble and insoluble impurities such as soil, protein, fats, waxes, gums, fine bagasse, and coloring matter in addition to soluble salts. The process designed to remove as much of these impurities as possible employs lime and heat and is called clarification or defecation. In simple lime defecation, milk of lime is added to the cold juice in amounts (about one pound of calcium oxide per ton of cane) sufficient to bring the pH to the range of 7.5 to 8.5. The limed juice is pumped through heaters in which the juice is heated to temperatures between 90 and 115°C. There are many modifications to this basic

Fig. 9.5. Lignex sugar cane separator, full frontal working view. (*Courtesy Lignex Products Group, Hawker Siddeley Canada Ltd.*)

A - Top-roller.
B - Bagasse roller.
C - Feed roller.
D - Trash plate.
E - Feed of cane.
F - Discharge of crushed cane.
G - Juice-collecting trough.

Fig. 9.6. A cane mill with its three crushing rolls and with a typical turn-plate arrangement. A hydraulic piston maintains pressure on A, while B and C are fixed. The rolls are under pressure of 75–80 tons per foot of roll width.

process involving different sequences of lime and heat. For example, fractional liming with double heating involves liming the cold juice to pH 6.4, heating, liming to pH 7.6, and heating again before sending the juice to the clarifiers.

The combination of lime and heat forms a flocculent-type precipitate with various components in the juice, which consist mostly of insoluble lime salts, coagulated protein, and entrapped colloidal and suspended matter. The precipitate is removed by sedimentation or settling in continuous closed-tray clarifiers (i.e., Rapi-Dorr, Graver, BMA, and Bach Poly-Cell). The juice leaving the clarifier is a clear brownish liquid.

The flocculent precipitate, or "muds," which settles on the clarifier trays, contains about 5 percent solid matter. Sugar is recovered from the muds by means of rotary vacuum filters equipped with perforated metallic screen cloth (Fig. 9.10). The turbid filtrate is returned to the clarification system, and the press cake is discarded or sent to the fields as fertilizer.

Recent innovations in the filtration of muds include the use of tighter filter media. A clearer filtrate is obtained, which can be sent to the evaporators without reprocessing. The Eimcobelt rotary vacuum filter uses a tight filter cloth that is continuously removed from the drum and washed with hot water. Another approach is the direct filtration of muds with horizontal leaf filters (Fas Flo Filters). One process uses polyelectrolytes to increase the filterability of the muds through cloth filters (Rapid-Floc system).

Good clarification depends upon the formation of a stable flocculent precipitate that settles rapidly. Juices low in phosphates and other minerals, such as magnesium salts, tend to be refractory. Fine soil carried in from the field also may interfere with clarification.[15,16]

Evaporation. The clarified juice (about 85% water) is pumped to evaporators where it is concentrated to a clear heavy syrup containing about 65 percent solids. Evaporation is carried out in multiple-effect evaporators in order to achieve maximum steam economy.

Fig. 9.7. Farrell four-unit milling train. (*Courtesy Farrel Co., Division of USM Corp.*)

Fig. 9.8. Continuous ring type total masceration sugarcane diffuser. (*Courtesy Sucatlan Engineering, Paris.*)

Fig. 9.9. Open top view of DDS double cylindrical tilted cane diffuser. (*Courtesy De Danske Sukkerfabrikker, Copenhagen.*)

Each effect is arranged in series and operated so that each succeeding one operates under a higher vacuum (lower pressure). This arrangement allows the juice to be drawn from one vessel to the next and permits it to boil at a low temperature. The concentrated juice (syrup) is removed from the last effect by a pump. Triple-, quadruple-, and quintuple-effect evaporators are used, but the quadruple-effect type is most common. In a four (quadruple)-effect evaporator, one pound of steam evaporates four pounds of water.

Crystallization. The syrup from the evaporator is pumped to a vacuum pan, in which it is evaporated to supersaturation in order to cause sugar to crystallize. The vacuum pan is a single-effect evaporator designed to handle viscous materials. It is a vertical cylinder with its bottom designed to allow easy removal of the crystallized mass. The heating elements used in vacuum pans either are short, large-diameter vertical tubes (calandria pans) or coil around the inner surface of the truncated cone of the pan (coil

Fig. 9.10. Oliver Campbell vacuum mud filter applied to cane juice muds. (*Courtesy Dorr-Oliver, Inc. Stamford, Conn.*)

pans). A typical vacuum pan with a "catchall" or entrainment separator for separating syrup from vapors measures approximately 25 feet in height and 15 feet in diameter. The working capacity is about 1000 cubic feet of massecuite (mixture of crystals and syrup or mother liquor). The shape of the pan and the positioning of the heating elements within the pan are important design factors in maintaining good circulation of the massecuite. For example, floating calandrias (calandria not attached to shell of pan), horizontal pans, and pans having mechanical circulators[17] are used.

The move toward continuous processing has resulted in the development of continuous vacuum pans. One review[18] of continuous vacuum pan technology provides considerable information on this important area. According to the author, pan boiling is one of the last remaining batch operations in sugar processing and has attracted considerable attention in recent years. As a result, "low grade continuous pan boiling is now well established in the South Africa sugar industry." It is stated that the benefits of continuous pan boiling include considerable thermal economy and steadier operation of the entire pan floor. Batch pans still are required to produce the seed for feeding to continuous pans, but, it is expected that this batch operation eventually will be replaced by a continuous process. It also is reported[19] that "continuous vacuum pans are now used in many beet factories (Europe) and cane factories (U.S., S. Africa, Australia, Brazil and others). They are also used in refineries in the U.S. and other countries mentioned above. Continuous crystallizers are in most European beet factories and many cane factories and refineries."

One version of the continuous vacuum pan is a horizontal cylinder with compartments in its lower part (Fig. 9.11). The pan is provided with an additional evaporator called the concentrator, where the density of the syrup is raised to 78 to 80 degrees Brix.* Seed is added in the first compartment, and the resulting massecuite moves progressively through the compartments of the pan. Additional syrup is added to each compartment to control the fluidity of the massecuite. Approximately 20 percent of the syrup is introduced at the concentrator, and 80 percent of the syrup is fed to the various compartments through special feed headers. The rate of massecuite discharge is controlled by the level in the pan.[20]

Fig. 9.11. Top view of open stork continuous crystallizer under construction. (*Courtesy Stork-Werkspoor Sugar, Hengelo, The Netherlands.*)

The crystallization of sucrose in vacuum pans is called "sugar boiling" and each boiling is termed a "strike." Because a single crystallization does not recover all of the sucrose from the syrup, mother liquor from a "strike" is recycled for recovery of additional sugar. There may be as few as two and as many as four reboilings to recover the maximum amount of sugar. A three-boiling system is the most popular. In order to distinguish among

*The Brix scale is a density scale for sugar (sucrose) solutions. The degrees Brix are numerically equal to the percentage of sucrose in the solution. The term "Brix solids" refers to the solids as determined by the Brix hydrometer.

the boilings, a letter is assigned to each boiling and its products; that is, the first boiling (syrup massecuite) would be given the letter *A*, and its products after centrifugation would be *A* sugar and *A* molasses; the second and third boilings would be assigned the letters *B* and *C*, respectively.

The boiling systems are based on "apparent purity," the amount of sucrose remaining in solution, and usually are expressed as the ratio of the polarization value to the total solids as measured by the Brix hydrometer. In a three-boiling system, the purity of the *A* massecuite is set between 80 and 85 by blending syrup with *A* molasses, the purity of the *B* massecuite is set between 70 and 75 by blending with syrup and *A* molasses, and the *C* massecuite is set between 55 and 60 by blending with syrup and *B* molasses. Other variations are possible.[6]

The *A* and *B* massecuites, after being discharged from a vacuum pan, are sent to centrifugals for separation. The *A* and *B* sugars are combined and become commercial raw sugar, the principal product of a cane sugar factory.

The *C* massecuite is a low-purity, highly viscous material that is not immediately sent to the centrifugals because of the large amount of recoverable sucrose remaining in solution. Instead, it is placed in crystallizers, U-shaped or horizontal containers equipped with coils attached to a hollow rotating shaft through which water circulates. The massecuite remains in the crystallizer from one to four days to allow additional crystallization to take place. Centrifugation of the *C* massecuite yields a final molasses (blackstrap) and a *C* sugar that is used for seeding the *A* and *B* boilings.

Centrifugation. Massecuite from the vacuum pans is sent to centrifugal machines wherein the crystals are separated from the mother liquor. A centrifugal consists of a cylindrical perforated basket lined with a screen of perforated sheet metal. The basket, enclosed in a metal casing, is mounted on a vertical shaft that rotates the basket, imparting centrifugal force to the massecuite.

In a batch centrifugal, hot massecuite is fed into the basket through a short chute from the holding vessel. As the basket rotates, the massecuite forms a vertical layer on the screen lining. When the machine reaches operating speed (1000–1800 rpm), the syrup flows through the perforated lining and basket and is removed through an outlet at the bottom of the casing. The sugar on the lining is washed with a spray of water to decrease the amount of molasses adhering to the crystals. The basket continues rotating until the sugar is fairly dry, at which time the machine is switched off and brakes are applied. The sugar is discharged by a manually or automatically operated plough.

In continuous centrifugals, the machines do not stop but continue in motion while receiving fresh supplies of massecuite. The rate of feed must be carefully regulated to obtain optimum separation of molasses from crystals (Fig. 9.12).[21]

Packaging and Warehousing. The packaging of raw sugar in jute bags has decreased to a very low level in most areas of the world. In fact, all of the raw sugar produced in Louisiana, Florida, Hawaii, Australia, and Puerto Rico is shipped in bulk. Where raw sugar still is bagged, bag weights vary from 110 to 150 lb in the Philippines to 250 lb in the West Indies. Bulk raw is brought from the factories to seaport terminals by means of dump trucks, railroad cars, or barges. These terminals, with capacities of 75,000 tons or more, are warehouses that were originally intended for bagged sugar, but now are specially designed structures for bulk handling. Unloading and loading of ships is done by bucket cranes or by gantries with traveling cranes.[13]

SUGAR REFINING

The raw sugar produced as described in the preceding section (called "raw" in the industry) is light brown in color and contains about 98 percent sucrose. In order for it to be competitive in the marketplace, it is necessary to bring the level of sucrose purity to 99.9 percent and to make it white. This is

Fig. 9.12. Silver continuous centrifuge. (*Courtesy CF&I Engineers, Inc., Denver, CO.*)

Fig. 9.13. Principal steps in manufacture of refined sugar. (*Adapted from Clarke, Margaret E., "Cane Sugar Process and Byproduct Streams," Iowa Biotechnology Consortium, March. 1991.*)

done in the refining process using the several steps depicted in Fig. 9.13.[8]

It may be helpful in establishing additional context for the previous section and for sugar refining to note that in 1970, there were four routes for the production of white sugar from sugarcane or raw cane sugar:[22] (1) traditional plantation white, obtained directly from sugar juice with settling clarification and sulfitation (SO_2 bleaching); (2) improved plantation white with carbonation and sulfitation; (3) refined white, from a white end refinery attached to a raw sugar factory, involving remelting (redissolving) raw sugar and refinery clarification and decolorizing processes; and (4) refined white, from a refinery not incorporated into a raw sugar factory.

Affination and Melting

The first step in the refining of raw sugar is called affination and consists of removing the film of molasses (in which a large portion of the impurities is contained) from the surface of the raw sugar crystal. The process involves the mixing of raw sugar with a saturated sugar syrup (72–75 degrees Brix) at about 120°F in a U-shaped trough called a mingler. (Saturated syrup is used to avoid dissolving the raw sugar.) The mingler has a rotating agitator to maintain maximum contact between the sugar and the syrup. The mixture is centrifuged to separate the crystals from the syrup, and the crystals are washed with hot water. The "washed" sugar is "melted" or dissolved in water to a density of 55 to 60 degrees Brix.

Potable water, steam condensate, and "sweet waters" are used for dissolution. The liquor from the melter is screened to remove nonsugar materials such as sand, stones, wood, cane fibers, and lint. Screening is done by vibrating screens (Tyler-Hummer), a cyclonic separator (Dorrclone), or centrifugal screens (DSM).[13,16]

Purification

A number of processes are available for purifying the liquor from the melter prior to its crystallization into granulated sugar. The two combinations most commonly used are:

1. Clarification–filtration–decolorization.
2. Clarification–filtration–decolorization–ion exchange.

Clarification. One of three types of chemical treatment is used as the first step in the purification process. These are liming, phosphatation, and carbonation.

Liming is the simplest of the chemical treatments. Process liquor is "limed" to a neutral pH using milk of lime and heated. A slurry of diatomaceous earth is added, and the treated liquor is pumped through a filter press, such as those discussed in the following section.

When phosphatation is used, the screened liquor is heated to 60 to 70°C and mixed with phosphoric acid (0.005–0.025% P_2O_5 on solids). The mixture is immediately limed to a pH of 7.0 to 8.0, aerated with compressed air, and sent to a clarifier (a rectangular tank equipped with heating coils). The liquor enters the clarifier at one end and is heated to 88°C while flowing to the outlet at the opposite end. The heated liquor releases air which rises, carrying the flocculent calcium phosphate precipitate to the surface. A blanket of this precipitate (scum) forms at the surface and is skimmed from the liquor surface by moving paddles. Clarifiers differ in shape and design; some of the best known are Williamson, Jacobs, Buckley-Dunton, and Sveen-Pederson.

In carbonation, process liquor heated to 60 to 80°C is limed to about pH 10 (0.03–0.8 percent CaO on solids), gassed with carbon dioxide, heated to 85°C, and regassed until the pH drops to between 8.4 and 9.0. Carbonation, because of the two-stage gassing, requires two clarifiers, primary and secondary. Washed flue gases are the source of carbon dioxide. The calcium carbonate precipitate that forms entraps colored matter, colloidals, and some inorganic compounds.[23]

A novel addition to phosphate clarification has been the Talofloc process for the cane refining industry and the Talodura process for the raw cane factory. The Talofloc process consists of the addition of two chemicals at specific points in the clarification station. First, a cationic surfactant called Talofloc is added to the unclarified melt liquor in concentrations ranging from 0.02 to 0.07 percent on solids. This step is followed by the addition of lime and phosphoric acid and the incorporation of air. Just before the liquor enters the clarifying chamber, a polyacrylamide flocculant (Taloflote) is added that induces the aggregation of the calcium phosphate precipitate, making a faster and more complete separation of the flow from the liquor. The Talofloc added prior to the chemical treatment causes at least a 50 percent increase in color removal when used in conjunction with Taloflote. The desugarization of the floc (scum) is best accomplished in a specially designed two-stage or three-stage countercurrent flotation system with further addition of Taloflote. For plants that are expanding or replacing their old clarifiers with the Talofloc process, special clarifiers designed to operate at maximum efficiency with Taloflote and Talofloc are recommended. One such unit with a capacity of 40 tons/hr is described in the literature.[24]

Filtration. The liquor from a phosphatation or a carbonation clarifier contains small amounts of finely dispersed particulate matter that require filtration for removal. The filtration process is similar for both types of clarified liquor.

A precoat of a filter aid (diatomaceous earth or perlite) first is deposited on the filter

surfaces of the press. (Additional filter aid is added to phosphate clarified liquors to improve press runs.) Liquor at 160 to 185°F is fed to the press at a pressure of about 60 psig until the flow rate drops below a predetermined level, ending the filtration cycle. The calcium carbonate particles in carbonated liquor act as a fairly good filter aid as long as the particles are developed to optimum size during carbonation. Large quantities of sugar polysaccharides (starch) in the melt liquor will restrict the growth of calcium carbonate, requiring the use of diatomaceous earth. A filter cycle may last from 2 to 12 hours depending upon the quality of the feed liquor.

Three types of filter presses generally are used, all of them leaf presses:

1. Horizontal body with vertical leaf filter, i.e., Sweetland, Vallez, Auto-Filters, and Industrial.
2. Vertical tank with vertical leaf filters; i.e., Niagara, Angola, and Pronto.
3. Vertical tank with horizontal plates; i.e., Sparkler and Fas-Flo.

The filter leaves are dressed with cloth made of cotton or synthetic fiber (nylon, dacron, etc.) or with a wire screen of 60 to 80 mesh size.

The sugar remaining in the filter cake is recovered by washing the cake in place (sweetening off) with hot water, or the cake is sluiced off the leaves with hot water and flushed to a holding tank for refiltration. The filtered sweet water generally is added back to the melter to dissolve washed raw sugar.[13,16]

Decolorization. Filtered clarified liquor is a clear, dark brown liquid having a solids content between 55 and 65 degrees Brix, a pH of 6.7 to 7.2, and temperatures between 65 and 85°C. If pulverized activated carbon is used as a decolorizing agent, it usually is added prior to filtration. Otherwise, the liquor is subjected to one or more decolorizing absorbents; for example, bone char, granular carbon, and ion exchange resins. Bone char removes colorants, colloidal matter, and some inorganic substances, whereas granular carbon is only a decolorizer. Ion exchange resins absorb color and change the composition of the ash.

Bone char and granular carbon are generally used in fixed beds or in cylindrical columns 20 to 25 feet high and about 10 feet in diameter. Some recent systems percolate liquor upward or downward through a stationary bed of adsorbent or by countercurrent flow of liquor and adsorbent (CAP or continuous adsorption process). (See below, Fig. 9.17.)

Liquor flow through bone char is about 1500 gal/hr over a period of 30 to 60 hours. For granular carbon, the flow rate is 3000 gal/hr for 20 to 30 days.

After the decolorizing cycle is completed, the adsorbent is sweetened off by displacing the liquor with water, and is washed. It is transferred to regenerating equipment consisting of dryers, kilns, and coolers. Bone char is regenerated at 540°C in a controlled amount of air. Granular carbon is revivified at 950°C in a limited oxygen and steam atmosphere. The kilns can be either the retort type on multiple hearths in columns (Herreshoff). After regeneration, the adsorbent is returned to the system for a new decoloring cycle.

Ion exchange resins are used in columns 8 to 10 feet high and 6 to 10 feet in diameter holding between 100 and 300 ft^3 of resin. The resin bed depth is only 2 to 4 feet. Flow rates are rapid (3000–4500 gal/hr), and the cycle is short (8–16 hours). Regeneration is accomplished in the column with a 10 to 15 percent salt solution at 140 to 160°F. The chloride form of a strong anionic resin decolorizes the liquor, and the sodium form of a strong cation resin softens the liquor. The sweet water from the various adsorbent columns are recovered by returning them to the melter.[13,25]

Crystallization

Decolorized liquor is a pale yellow liquid with a solids content of 55 to 65 degrees Brix. In some refineries, this liquor goes directly to a vacuum pan for crystallization, but most refineries (both cane and beet) preevaporate

the liquor in a multiple-effect evaporator to a solids content of about 72 degrees Brix.

There are four stages in the process for crystallization of sucrose: (1) seeding or graining; (2) establishing the seed; (3) growth of crystals; and (4) concentration.

A sufficient quantity of evaporated liquor is drawn into the pan to cover the heating elements. Water is evaporated from the syrup until its supersaturation approaches 1.25. At this point, the steam pressure is lowered, and seed crystals are added. The seed is finely pulverized sugar dispersed in isopropyl alcohol or sugar liquor. This method is called shock seeding because addition of the seed induces an immediate formation of crystal nuclei throughout the supersaturated syrup. The nuclei are grown to a predetermined size or grain. Once the grain is established, the crystals are grown to size by maintaining supersaturation between 1.25 and 1.40 through control of steam pressure, vacuum, and the feed rate of the evaporated liquor. Adequate circulation during crystal growth is important.

When the volume of the massecuite reaches the maximum working capacity of the pan, the syrup feed is shut off, and evaporation is allowed to proceed until a thick massecuite is formed. When the massecuite concentration is considered "tight," the steam and the vacuum are shut off, and the massecuite is dropped into a holding tank equipped with agitators, where it is kept in motion until discharged into centrifuges.

Instrumentation is used extensively in sugar boiling to control the process of crystallization. Some of the principles used to provide control of vacuum pans for boiling a strike of sugar are:

1. Boiling point rise (BPR): thermometers are used to measure the temperature of the massecuite and its vapors (difference between these two temperatures is the observed BPR).
2. Electrical conductivity of the massecuite, based on the principle that conductivity is inversely proportional to the viscosity of the solution, which in turn has a similar relationship to the water content and thus the degree of supersaturation.
3. Fluidity of the massecuite: an ammeter is used to measure the current used by the motor of a mechanical circulator; changes in current indicate changes in the fluidity of the massecuite.
4. Soluble solids, measured by a refractometer.

Several systems employing instrumentation for automatic control of pan boilings by supersaturation have been used successfully (Fig. 9.14). However, controls are empirical in nature, and are based on the conversion of the various techniques and arts of a "sugar boiler" into mechanical operations that lend themselves to automation.

The boiling system of a refinery is straightforward: the first strike is boiled from evaporated liquor, the second strike is boiled from the runoff syrup of the first strike, continuing on to three or four strikes. The runoff of the last strike can be used in a variety of ways: (1) as syrup for affination, (2) as syrup for remelt, (3) or as syrup for producing a soft brown sugar; or (4) it can be reprocessed and reboiled.[13,17]

An innovative approach to the conversion of two-batch vacuum pans into a continuous crystallizing system was reported at the 1979 meeting of Sugar Industry Technologists.[26]

Remelt Sugars

The sugar contained in affination syrup must be recovered for economy. Recovery is accomplished by crystallizing the sugar in a vacuum pan. The resulting sugar is "raw" in composition and is returned to the refinery at the affination station. The residual syrup from the remelt station is known as refiners' blackstrap.

Centrifuging

Refined sugar crystals are recovered from the mother liquor by centrifuging the massecuite in equipment similar to that used for affining raw sugar. However, at this point, the crystals are washed with a greater quantity of hot

Fig. 9.14. Schematic drawing of vacuum pans and controls. (*Courtesy the Foxboro Co. Reproduced from* Sugar y Azucar, *Oct. 1970.*)

water. The washed crystals are discharged into a holding bin supplying a dryer.

Drying

Sugar from the centrifugals contains 1 to 2 percent moisture and is too wet to be placed in storage or packages. The wet sugar is fed to drying equipment called granulators, which usually are rotating drums, 15 to 35 feet long and 6 to 7 feet in diameter, inclined slightly so as to be discharged by gravity. Heated air is blown through the dryer concurrent with the flow of the sugar. Two types commonly used are the Roto-Louvre granulator and the Standard-Hershey granulator. The Roto-Louvre is a single rotating drum in which hot air from louvres in the wall dried a moving bed of sugar. The Hershey granulator consists of a rotary dryer followed by a second unit in which the sugar is cooled to 45 to 55°C after leaving the dryer at 52 to 55°C.

Packaging and Storing of Refined Granulated Sugar

In recent decades, there have been significant changes in the methods of handling refined sugar as it leaves the dryer. Bulk delivery of refined granulated sugar to customers has made it necessary for the refiner to store finished product in bulk rather than in bags. Granulated sugar remains free-flowing for a longer period of time if it is "conditioned," as conditioning reduces handling, packaging, and storage problems resulting from caking. Conditioning involves several factors, including control of the moisture content, the temperature, and the grain size, as well as good inventory management.[27]

As the sugar leaves the dryer, it can be cooled in one of several types of equipment, including the Holoflite screw conveyor, the thermal disc processor, and the Buttner turbo tray dryer; or it can proceed directly to large bins or silos, where it is cooled and conditioned in an atmosphere of controlled temperature and humidity. Circulation of dehumidified air through or above the sugar pile and mechanical movement from silo to silo also are practiced to prevent caking. The cooled, unconditioned refined sugar then is classified according to crystal size by a screening or sieving operation.

Shipment of conditioned bulk sugar is made in specially designed railcars (100,000–180,000 lb capacity) or in hopper trucks (40,000–60,000 lb capacity).

Conditioned sugar also is packaged in

multiwall paper bags of 2-, 5-, and 10-lb capacity. These sizes are overwrapped in 60-lb bales for shipment. Larger bags (100, 50, and 25 lb) also are packed.

Granulated sugar is available in several sizes according to the needs of customers. These range from very large to very small crystals.

Direct-Consumption Sugar

Some cane sugar mills produce white sugar by using sulfitation, carbonation, or, more recently, more complex processes such as Blanco Directo.

Sulfitation. The expressed juice from the mills is heated to 75°C, clarified with lime and sulfur dioxide, and in some cases filtered and then evaporated. After the syrup is treated with sulfur dioxide again, it goes through a three- or four-boiling system. The *A* and *B* sugars are mixed with a high-purity syrup, and then centrifuged, dried, and screened for size and distributed as white sugar.

Blanco Directo. Talo "Blanco Directo" is a trademark of Tate and Lyle Process Technology. This process and others (e.g., USP process of the Fabcon Co.) are used at sugarcane factories, in conjunction with raw or plantation white production, to produce a high grade white sugar suitable for use in food and beverages. This product is economically competitive with imported refined sugars. Although the term "Blanco Directo" is the trademark for one particular process, it appears to be used somewhat generically for many similar processes.

In using Blanco Directo or similar processes, instead of shipment of the raw sugar from a sugar factory to a refinery to be refined into white and brown sugar, the refined product is manufactured at the factory via the following steps:[28]

1. Phospho-flotation of juice or syrup to remove turbidity, ash, and color.
2. Addition of flocculation aids and ionic organic polymers to aid in color removal.
3. Flotation clarification of mud filtrate to remove nonsugars.
4. Sulfitation (but to a much lesser degree than in traditional plantation white) after initial partial liming of the juice.

According to Clarke.[72] "An important step in the process is clarification of the syrup coming from the evaporators, to reduce turbidity, color, and ash. Syrup clarification, such as the Talodura process, has in this decade (1980s) found widespread application in raw sugar production in the U.S., throughout the Far East, and Central and South America." Talodura consists of using a combination of sulfite clarification followed by subsequent clarification of the evaporated liquor with lime and phosphoric acid to which a flocculant is added. It has made improvements in the color of the mill white sugar, boiling-house recovery, and chemical costs per ton of sugar.[29]

Carbonation. In this process, the mixed juices from the mills are heated, clarified with lime, and evaporated to about 35 degrees Brix. This syrup is relimed and treated with carbon dioxide, filtered, recarbonated, reheated, and refiltered. After carbonation, the syrup is given a double sulfur dioxide treatment and filtered. The resulting syrup is subjected to a three- or four-boiling system with the *A* and *B* sugars used as the refined product.

Other Sugars

The cane and beet sugar industries manufacture several types of sugar in addition to granulated sugar. Among them are liquid sugars, brown or soft sugar, pulverized sugar, and special agglomerated sugars.

Liquid Sugars (Sugars Dissolved in Water). Liquid sugars are economically important largely because a number of food manufacturers use sugar in the form of a syrup, and also because of the efficiency and ease of handling of a liquid product. Liquid sugars are produced by one of two methods: (1) dissolution or melting of refined sugar in

water; (2) further purification of in-process liquors by ion exchange treatment to remove minerals, and further decolorization using bone char and pulverized or granular carbon. Liquid sugar had other advantages over granulated in that it can be delivered as sucrose, as a mixture of sucrose and invert sugar, or as a blend of those sugars with various glucose syrups from starch hydrolysis and dextrose from the same source. Sucrose liquid sugars usually are distributed at a 67 percent sugar concentration, whereas invert sugar and mixtures of invert and sucrose are distributed at concentrations of 72 to 77 percent sugar. Invert sugar is an equimolecular mixture of dextrose and levulose resulting from the hydrolysis of sucrose.

Brown or Soft Sugars. These are specialty "dry" products ranging in color from light to dark brown. In addition to sucrose, soft sugars contain varying amounts of water, invert sugar, and nonsugars. These sugars have a characteristic flavor. In some cases, soft sugars are "boiled" from a low-purity process liquor to obtain the desired color, flavor, and composition. In other instances, refined granulated sugar is "painted" with refiner's syrup or molasses to produce a product of similar appearance and characteristics.

Pulverized Sugars. These sugars are manufactured by milling granulated sugar to the desired size. In most instances, pulverized sugars are mixed with small amounts of dried starch or tricalcium phosphate to prevent caking. A free-flowing agglomerated pulverized sugar without anticaking agents has been manufactured.[24]

Agglomerated Sugars. Several kinds of these sugars are manufactured. Because soft sugars are subject to severe caking, one sugar refiner has resolved the problem by converting soft sugar into an agglomerated free-flowing dry product. Several companies manufacture "dry fondant" by agglomerating mixtures of invert sugar and very fine (less than 40 microns) sucrose particles.

Amorphous Sugar. Amorphous sugar is a microcrystalline sugar agglomerated during crystallization, similar to dry fondants without having invert or reducing sugars added to it. Its polarization ranges from 99.1 to 99.2 with an invert content of from 0.01 percent to 0.15 percent, a color of 20 to 40 m.a.u., and a size in the 40 to 80 mesh range.[30]

BEET SUGAR

Agriculture

The sugar beet, *Beta vulgaris*, is a Temperate Zone plant grown largely in the Northern Hemisphere. A long growing season is required, and in the United States it is one of the first crops planted and the last harvested. The seed is planted in rows, and after the plants have emerged, they are "thinned" to permit better development of a smaller number of beets. Unlike sugarcane, which cannot be cultivated after the cane has reached a certain height, sugar beets require a considerable amount of cultivation. Both cultivating and harvesting operations have been largely mechanized in the United States.

After the plants have reached maturity in late fall (October and November), they are harvested by machines that remove the top growth of leaves, lift the roots from the ground, and deliver them to a holding bin or a truck. Because the harvesting season is much shorter than the processing season, beets are stored in piles at the factory or at outlying points near transportation. Frequently, storage piles are ventilated to lower the temperature of the beets, thus reducing sugar loss due to respiration during storage.[7,31]

In the Red River Valley area of Minnesota and North Dakota, storage sheds for sugar beets were constructed in the mid-1980s that measure 210 feet wide by 520 feet long and hold 50,000 tons of beets. The temperature in these storage facilities is controlled at not much above freezing to minimize degradation of the sucrose content.

Washing

Beets are transferred from storage into the factory in water flumes. These flumes lead directly into a rock-catcher, which allows rocks to settle out, and then on into a trash-catcher and a rotary washer.

Slicing

The washed beets are sliced into thin V-shaped cossettes by means of specially shaped knives set in frames mounted around the periphery of a rotating drum. Good removal of rocks and trash is essential to a reasonable life for the knives.

Diffusion

The cossettes are weighed and transferred continuously into a diffuser where water passes countercurrently to the movement of the cossettes, and where, by dialysis, sugar and some of the nonsugars of the beet are extracted.

Both continuous and batch diffusers are used, but batch processes largely are being displaced. Continuous diffusers come in a variety of forms and shapes, but all of them employ the same principle, that is, movement of juice countercurrently to the movement of the cossettes. In batch diffusers, the cossettes are held stationary, and the diffusion juice is moved from cell to cell. A few of the well-known types of continuous diffusers are the RT, BMA, DDS, DeSmet, Buckau-Wolf, and Olier.

The RT-type diffuser is a large revolving drum with an internal helix that separates the drum into moving compartments. As the drum revolves, the cossettes travel by the action of the moving helix from one end of the drum to the other end, while the juice moves in the opposite direction.

The BMA diffuser is a cylindrical tower with a conveyor mechanism attached to a central rotating shaft. Guide plates on the shaft direct the cossettes upward, while the juice exits through screens at the bottom.

The temperature of the juices during diffusion ranges up to 60 to 75°C. Antifoaming agents are used to control foam in continuous diffusers, and bactericides are added for microbiological control.[7]

Juice Purification

The diffusion juice is a dark turbid liquid containing 10 to 15 percent sucrose and 1 to 3 percent nonsugars, including proteins, nitrogenous bases, amino acids, amides, inorganic material, and pectinous matter. It also contains $\frac{1}{4}$ percent sodium and potassium. These impurities are removed by a series of processes using lime and carbon dioxide (carbonation), sulfur dioxide (sulfitation), and crystallization. In recent years, there has been an increasing use of ion exchange resins in these plants where refined granulated sugar is made.

Carbonation. In this process, the juice is heated to 80 to 90°C, limed, and then sent to a carbonator for gassing with carbon dioxide. The resulting mixture, containing insoluble lime salts, chiefly calcium carbonate, is pumped to subsiders (thickeners, clarifiers) to remove the insolubles by settling. From the subsider, the partially clarified filtrate is recarbonated, and the residual lime is precipitated. This treatment is followed by press filtration. The sludge from the subsiders is filtered on rotary-drum filters, and the sugar-laden filtrate is returned to the first carbonation step. The carbonation process can be either continuous or batch.

One continuous system for carbonation used in the United States is the Dorr process, introduced in the 1930s. In this process, heated juice is treated with milk of lime or lime saccharate in a cylindrical column (primary tank) where carbon dioxide gas is added. The contents of the secondary tank are continuously pumped into the bottom of the primary tank. The finished first-carbonated juice flows from an overflow line in the discharge of the recirculation pump to a subsider in which the insoluble matter is separated from the juice. The thickener is a cylindrical tank consisting of a number of shallow compartments with trays and rotating paddles that continuously remove the settled sludge. The filtrate is

reheated (90–95°C) and sent to another carbonation tank. The effluent from the second carbonation tank is pressure-filtered on a leaf press such as those manufactured by U.S. Filter, Industrial Filter, Sweetland, Vallez, or Angola. (Membrane filters are beginning to replace presses.) The filtered liquor is treated with a small amount of sulfur dioxide in a tower, heated, and refiltered.

Pre-liming. The Dorr process is being supplanted in the United States and Canada by European technology that uses a pre-liming system. In pre-liming, diffusion juice is sent through a series of tanks where lime is added to various levels of alkalinity. The limed juice then is carbonated and generally follows a purification pattern not unlike that described for the Dorr process. The pre-liming system allows the process to handle a wider range of beet quality and leads to a more uniform product. According to one authority,[71] the DeDanske Sukkerfabrikker integrated juice purification system, which incorporates fractional cold liming, increases the yield of high-quality white sugar while effecting marked cost reductions.

Ion exchange resins are used in some plants to decolorize, deionize, and delime the juices.

Evaporation and Standard Liquor

The thin juice (about 15% solids) discharged from the filter presses after clarification is evaporated to thick juice (50–65% solids) in multiple-effect evaporators. If decolorizing adsorbents are used, they usually are added to the thick juice. Granular carbon and pulverized carbon have been used for this purpose. Thick juice treated with pulverized carbon usually must be double-filtered to remove all of the carbon.

Low raw sugars (sugars with purities below that of refined) are added to the thick juice in the melters to make standard liquor from which white sugar is crystallized. Standard liquor usually is filtered with diatomaceous earth before going to the vacuum pan.

Crystallization, Centrifuging, and Drying

Crystallization practices in U.S. beet sugar factories are similar to those in a cane sugar refinery and result in a white granulated sugar comparable in quality to refined cane sugar. However, some European factories may make "raw sugar" as a separate product. White-sugar centrifuge stations as well as drying operations are comparable to the same operations in a cane sugar refinery.

Sugar Recovery from Molasses

Molasses is the material that remains when no more sugar can be economically extracted in the crystallization process. Generally it is quite high in sugar content (60% dry basis) and represents considerable value. It is now possible, using recently developed ion exclusion chromatography, significantly to increase the recovery of these sugar values.

Historically, recovery of sucrose from molasses was accomplished by using the Steffan's process, in which sucrose is separated by means of a lime salt known as calcium saccharate. Another process, known as the barium process, precipitated sucrose in the form of barium saccharate. Descriptions of both of these processes can be found in the preceding edition of this book and elsewhere. Because they both are being replaced by ion exclusion, only that process is described here.

The ion exclusion process utilizes a bed of ion exchange resin operated in such a way as to simulate a countercurrent moving bed system. Its rationale and theory have been discussed in detail in at least one paper.[32] The ion exclusion process for desugarization of molasses provides for lower sucrose content in the final molasses—up to 90 percent of the sucrose in the molasses can be recovered—and has been put into every major beet sugar operation in the United States. Suppliers of the technology are Amalgamated Sugar, Finnish Sugar, and Illinois Water Treatment.

The resin used as a strong-acid cationic resin, crosslinked with divinylbenzene to form a polymer network, in the monovalent salt form. It is placed in a column alternately fed

with molasses and water. The height-to-diameter ratio of the column is important (not less than 4 to 1 and preferably around 20 to 1), and flow rates of from 0.4 to 0.6 gal/ft^2/min have been suggested. The clarity and temperature (70–80°C) of the feed also are important. Sugar and nonsugar components of the molasses are separated chromatographically. Four fractions are obtained: the first fraction contains most of the inorganic substances, colorants, and organic nonsugars; the second fraction is similar to the first in composition of the feed but contains about 25 percent more water; the third fraction is a high-purity sucrose solution between 10 and 20 degrees Brix; the fourth fraction is an extremely dilute solution of sucrose and invert sugar (2–5 degrees Brix). The sucrose fraction either goes to pan crystallization or is further purified with activated carbon and ion exchange resins and sold as liquid sugar. The first or waste fraction is concentrated to 70 degrees Brix and marketed as "fodder molasses." The resin, after a number of cycles, has to be regenerated with salt because of the accumulation of calcium and magnesium ions. In some cases, the feed molasses first is softened prior to ion exclusion treatment.[33,34]

Sugarcane molasses can be subjected to the same basic ion exclusion process with some modification: (1) the molasses must be clarified to remove as much suspended matter as possible, and (2) prior softening is necessary to extend operating cycles. Because cane molasses contains substantial quantities of reducing sugar, the resulting fractions consist of nonsugars, a mixture similar in composition to feed, a mixture of sucrose and reducing sugars rich in sucrose, and a mixture of reducing sugars and sucrose.[35,36]

Sugar Recovery from Syrup

In 1962, a beet sugar company in the United States built a tank system for storing thick juice produced in excess of its crystallizing capacity. This storage system served two purposes: (1) it enabled the "beet end" of the factory to operate at full capacity even though the "sugar end" could not handle the output; and (2) it provided raw material for the sugar end to operate at full capacity when the beet end was in trouble (Fig. 9.15).

Fig. 9.15. Schematic flow diagram showing flow of thick juice to storage. (*Courtesy the Amalgamated Sugar Co., Ogden Utah.*)

Fig. 9.16. Flow diagram showing flow of thick juice from storage through processing. (*Courtesy the Amalgamated Sugar Co., Ogden, Utah.*)

The system proved such a success that other beet factories adopted the practice. Furthermore, improvements made in the storage systems have made it possible to extend the operating time of the sugar end far beyond the termination of the slicing compaign (Fig. 9.16).[37]

Storage and Packaging

In the past several years, both beet sugar factories and cane sugar refineries have gone extensively to bulk storage of finished products. Beet sugar bulk storage usually is more extensive because of the need to hold sugar to meet customer needs from one processing season to the next. The daily capacities of cane sugar refineries are such, however, that it would be uneconomical to store more than a few days' output.

Prior to bulk storage, refined beet sugar was stored in bags (mostly 100 lb). Such storage was acceptable as long as large customers purchased sugar in bags. However, a shift in the consumption of sugar from the home to food processors and the development of bulk transportation facilities made bulk storage a natural consequence. Thus, beet sugar is packed in the usual customer-size packages and 100-lb bags, and it is also shipped in bulk containers ranging in size from "tote bins" (2000 lb) to railcars holding several tons.

Cane and Beet Sugar Production

Cane sugar is primarily a product of tropical countries, whereas beet sugar is produced in temperate climates. However, both are produced in a few countries, among them the United States, China, Afghanistan, Spain, Iran, and Pakistan.[38] The total annual production of centrifugal or crystalline sugar is shown in Table 9.1, together with average annual production for representative periods over the years 1985–90.

Cane sugar production currently amounts to approximately 65 percent of the total world sugar production. The largest seven producers are EC-12, India, the former Soviet Union, Cuba, Brazil, China, and the United States, in that order.[39]

The United States in 1977–78 produced less than 6 million short tons and in 1989–90 6.6 million short tons, raw value, of sugar.[75,76] About half of this amount was cane sugar produced in Florida, Hawaii, Louisiana, Texas, and Puerto Rico. The remainder was beet sugar produced in all the states west of the Mississippi River except Arkansas, Missouri, Oklahoma, and Nevada, and, in addition, in the states of Michigan and Ohio.

Sugar Consumption and Usage

Per capita consumption of sugar varies widely from country to country. Although Europe, the Americas, and Oceania consume more than 70 kg per capita, the average for the world is only slightly more than 37 kg.

TABLE 9.1 Centrifugal[a] Sugar Production Raw Value[b] by Continent, 1986/86–1989/90 (in Units of One-Million Metric Tons)

	1985/86	*1986/87*	*1987/88*	*1988/89*	*1989/90*
North America	19.9	20.9	20.9	21.0	20.4
South America[c]	13.0	13.6	13.3	13.7	12.8
Europe (west)	16.3	16.7	15.8	16.4	17.2
Europe (east)	4.8	4.8	4.6	4.2	4.8
Africa	7.7	7.9	7.7	7.7	7.8
Oceania	3.8	4.0	4.0	4.1	4.3
U.S.S.R (Europe & Asia)	8.3	8.7	9.6	8.9	9.5
World Total	98.7	103.5	103.4	105.5	108.2

[a]Noncentrifugal sugars are usually consumed in the area of origin. Production from 1979 to 1980 ranged from 10.4 to 11.9 million metric tons (1979/80).
[b]Raw value is a term defined mathematically to reduce sugars of differing sucrose contents (from raw sugar containing varying amounts of sucrose to refined sugar containing approximately 100% sucrose) to an equivalent basis in pounds.
[c]Includes the Caribbean and Central America.
Source: U.S. Department of Agriculture.[75]

Sugar usage patterns, particularly in developed countries, have not changed markedly over the years except in the United States, where high fructose corn syrup (HFCS) replaced sugar in the manufacture of beverages (see Table 9.2). Nonfood uses continue to be a very small portion of the total.[73]

Although the nonfood uses of sugar are relatively small in comparison with its food use, because of its availability in large quantities at very high purities sugar always has been viewed as an attractive raw material for organic chemical synthesis. Research in this area has been aggressively pursued for more then 25 years,[40] and a great number of products have resulted from those efforts although few of them have been commercialized, either for economic reasons or because of the complexity of the chemical synthesis. Among the derivatives produced in the laboratory over the years are ethers, fatty and other esters, acetals, reduction products, halogen, sulfur, metal derivatives, and others.[41] Potential applications include products for use such as surfactants, surface coatings, food and feed additives, polymers, textile chemicals, pharmaceuticals, and pesticides.[42]

The reduction products sorbitol and mannitol have been manufactured from sugar and sugar products. Sugar's use for that purpose usually depends upon its being competitively priced with dextrose or glucose syrups, which are the most commonly used raw materials. A review of the potential applications of sucrose derivatives in other areas is contained in *Sucrochemistry*, ACS Symposium Series 41, 1977, American Chemical Society, Washington, DC.

Fatty acid esters of sucrose have been manufactured in the United States and some foreign countries on a commercial basis for many years.[40] Plants for a solventless process for manufacturing esters are in various stages of design and construction. An almost ironic laboratory development, casting chlorosucrose derivatives in the role of a noncaloric sweetener, has received widespread interest in recent years.[43,44]

TABLE 9.2 U.S. Sugar Deliveries to Industrial and Nonindustrial Users, Calendar Years, 1978–90 (Biennial)

	Calendar Year						
Type of Buyer	*1978*	*1980*	*1982*	*1984*	*1986*	*1988*	*1990*
	1000 short tons, refined						
Industrial users (food use):							
Bakery and cereal	1265	1284	1296	1404	1432	1541	1607
Confectionery products	895	895	939	1115	1051	1107	1277
Processed foods	672	516	1150	433	387	354	331
Dairy products	538	436	404	408	447	394	460
Beverage use	2558	2119	1583	908	266	237	228
Other	407	548	526	416	443	529	641
Total industrial use	6335	5798	5898	4684	4026	4162	4544
Nonindustrial Users:							
Institutions	104	96	85	108	84	89	106
Wholesale grocers	1963	1841	1951	1744	1867	2200	2121
Retail grocers	1168	1150	1086	1100	1066	940	1071
Others	124	205	92	101	58	86	75
Total nonindustrial use	3359	3292	3214	3053	3075	3315	3373
Nonfood use	158	117	106	127	138	121	107
Total use in U.S.A.	9852	9207	9218	7864	7239	7598	8024

SUGARS DERIVED FROM STARCH

Starch Hydrolysis

Undried starch, usually from wet milling, is pumped as a slurry to the conversion plant where it undergoes one or more hydrolytic processes to yield mixtures of various carbohydrates in the form of syrups or as crystalline dextrose. The kind and amount of the various carbohydrates obtained depend upon the type of hydrolysis system used (acid, acid–enzyme, or enzyme–enzyme), the extent to which the hydrolytic reaction is allowed to proceed, and the type of enzyme used. The fact that most starches consist of two different kinds of polymers (amylose and amylopectin) also has an effect on the nature of the products obtained.

The extent to which starch is converted into simpler carbohydrates is indicated by the dextrose equivalent (D.E.), a measure of the reducing sugar content calculated as dextrose and expressed as a percentage of the dry substances. Hydrolyzates having dextrose equivalents ranging from 5 to 100 are produced. Those having a very low dextrose equivalent frequently are referred to as dextrins. They are produced by minimal acid hydrolysis or roasting. Starch-hydrolyzate syrups commonly are produced as "low," "regular," "intermediate," "high," or "extra-high" conversion products, as more or less standard products.

Table 9.3 shows the composition of some of the various syrups. In addition to carbohydrates, the syrups contain minerals (largely sodium chloride) and nitrogenous substances. Maltodextrin is a dried dextrose having a D.E. below 20, while the dry dextrins with a D.E. above 20 are called corn solids.

Acid Hydrolysis. A starch slurry containing 35 to 45 percent solids is acidified with hydrochloric acid to about pH 1.8 to 1.9. The suspension is pumped into an autoclave (convertor) where live steam is gradually admitted to a pressure of 30 to 45 psi. The conversion time largely determines the D.E. of the hydrolyzate; e.g., 8 minutes may produce 42 D.E. syrup, and 10 minutes may produce 55 D.E.[45] Converted liquors are neutralized with sodium carbonate to a pH of 5 to 7, with coagulation of insoluble protein, fats, and colloidal matter. The scum is removed by centrifuge.

The dark-colored clarified liquor is pressure-filtered and then concentrated to 60 percent solids in multiple-effect evaporators. The concentrated liquor is decolorized with granular carbon in columns 12 feet in diameter and 30 feet high in a countercurrent manner; that is, liquor flows upward in the columns, while a portion of the carbon is removed from the bottom periodically (Fig. 9.17). It has been reported that carbon is used at a rate of 2.5 percent of dry solid processed, and that 5 percent of carbon is lost during revivification.[46] "Low ash" syrups usually are deionized with ion exchange resins. Finally, the processed liquor is evaporated to a final solids content of 75 to 85 percent in a single-effect evaporator.

TABLE 9.3 Carbohydrate Composition of Glucose Syrups (Saccharide as a % of Total Carbohydrates)

DE	*Mono*	*Di*	*Tri*	*Tetra*	*Penta*	*Hexa*	*Hepta*	*Higher*
15	3.7	4.4	4.4	4.5	4.3	3.3	3.0	72.4
35	13.4	11.3	10.0	9.1	7.8	6.5	5.5	36.4
45	21.0	14.9	12.2	10.1	8.4	6.5	5.6	21.3
55	30.8	18.1	13.2	9.5	7.2	5.1	4.2	11.9
65	42.5	20.9	12.7	7.5	5.1	3.6	2.2	5.5

Source: Corn Industries Research Foundation, "Corn Syrups and Sugars," Washington, DC, 1965.[70]

Fig. 9.17. Flow diagram of Pittsburgh moving bed system for glucose-dextrose. (*Courtesy Activated Carbon Div., Calgon Corp.*)

Acid–Enzyme Hydrolysis. Starch first is liquefied and hydrolyzed to specific dextrose equivalents with hydrochloric acid. After evaporation to 60 percent solids, a saccharifying enzyme (fungal amylases) is added to continue hydrolysis to the desired level. By choosing two or more types of enzymes (such as alpha-amylases, beta-amylases, or glucoamylases) and adjusting the initial acid hydrolysis, syrups with different ratios of dextrose, maltose, and higher saccharides can be obtained.[47]

Enzyme–Enzyme Hydrolysis. Enzyme–enzyme conversion employs an enzyme for starch liquefaction in place of acid. Subsequent hydrolysis is by enzymes as above. The choice of hydrolytic system depends upon economics and the kind of end product desired. Enzymes usually are inactivated by heating to 75 to 80°C.[47]

Conversion to Dextrose

The manufacture of dextrose requires conditions different from those used for the manufacture of syrups. When acid hydrolysis is used, the density of the starch slurry is lower

(15–20%). Hydrochloric acid is added to a concentration of 0.03 N, and higher steam pressures (40–50 psi) are used. The time of hydrolysis is extended to about 30 minutes or until the purity reaches 90 to 91 D.E. The hydrolyzate is neutralized to a pH of 4 to 5 with sodium carbonate. Other purification steps are similar to those employed in syrup manufacture.

The high-purity liquor is pumped to insulated crystallizers fitted with slowly moving agitators for crystallization of dextrose monohydrate. A heavy seed base (about 25%) from a previous batch is mingled with the syrup and cooled to about 38°C. The seeded liquor is held at this temperature for several days until about 60 percent has crystallized. The mixture then is centrifuged to separate crystals from the mother liquor. The wet sugar is dried in rotary dryers or recrystallized into anhydrous dextrose. The monohydrate also may be converted to anhydrous dextrose by drying the monohydrate in hot air. A second crop of crystals is taken from the mother liquor, and the runoff syrup from this step is final molasses or "hydrol."[48,49]

Dextrose also may be made by acid–enzyme and enzyme–enzyme systems. In the acid–enzyme process, acid hydrolysis is stopped near 18 D.E. The liquefied slurry is evaporated to 60 percent solids and treated with a fungal glucoamylase until the D.E. is about 90. With the enzyme–enzyme process, a 30 percent starch slurry is liquefied with a heat-stable bacterial amylase followed by treatment with fungal glycoamylase. The liquor is purified in the same manner as used for others.[47]

In addition to syrups, several solid products are manufactured, including several kinds of dried sugar. "Crude sugar" is a solidified liquor containing large amounts of dextrose and substantial quantities of nondextrose solids as well. A solid product of recent development, "noncrystalline" or "enzyme-converted" dextrose, is highly refined and contains about 92 percent dextrose.

High Fructose Syrup

The rapid development of enzyme technology has made available in commercial quantities a new sweetener derived from starch. This new product, which is generally known as high fructose corn syrup (HFCS) contains 30 percent fructose, 35 percent glucose, and 6 percent higher saccharides. The fructose fraction is obtained by the isomerization of glucose through the catalytic action of the enzyme glucose isomerase.

Glucose isomerase can be derived from a variety of microorganisms; *Bacillus coagulans*,[50] *Actinoplanes missourensis*,[51] and several organisms from the *Streptomyces* genus.[52] Because glucose isomerase is an intracellular enzyme, it commonly is used today in an immobilized form fixed on an inert carrier, geometrically shaped for use in vertical columns.

The basic feature of the isomerization process is a series of reactors containing the immobilized enzyme fixed in a packed bed. The feed material is of the highest practical purity possible because enzymatic activity and life are closely related to the purity of the glucose syrup. Glucose syrup of about 94 D.E. is filtered, treated with activated carbon to remove residual color, and deionized with ion exchange resin to lower the ash content, particularly the calcium ion. The dissolved oxygen is reduced by flash evaporation, which also concentrates the feed stock to 40 to 45 percent glucose and raises the temperature to 60 to 65°C. At this point, prior to entering the reactors, enzyme activators are added, particularly magnesium ions and, in some cases, hydrogen sulfite ions. Also the pH is adjusted to about 8. Isomerized liquor is removed from the process when the equivalent of 42 percent fructose on a dry basis is reached. The isomerized syrup is further treated with activated carbon and ion exchange resin and finally concentrated to 71 percent solids.[46,50,54]

In the late seventies, a "second generation" of HFCS became commercially available containing 55 percent fructose on a dry basis. The higher fructose content is achieved by subjecting the 42 percent fructose syrup to a chromatographic separation of fructose from glucose by means of strong cationic resins of

low cross-linkage in the calcium form. The fructose is preferentially held by the resin and is later eluted with water.[46,55] Recently, a process has been described that uses a nonresin absorbent—a rigid, nonswelling inorganic product. The process employs a unique cycling device to remove from, or feed to, various parts of the chromatographic column liquids of different composition. The product stream has a fructose content of 95 percent on a dry basis.[56]

Fructose

Fructose, also known as levulose and fruit sugar, has been recognized as a naturally occurring sugar found in several fruits as well as honey for many years; but it has become commercially significant only in the past several years.[57] Fructose, as a laboratory chemical, has been prepared from inulin and invert sugar,[58] but its commercialization was based on its extraction from hydrolized sucrose solutions or invert sugar syrup (crystalline fructose).

Fructose has had wider usage in Europe, where it is manufactured by several companies. Those processes have been based on the hydrolysis of sucrose followed by chromatographic or ion exclusion separation, with subsequent crystallization of fructose after its separation from glucose and sucrose. The development of processes for isomerizing glucose into fructose, together with commercialization of chromatographic separation of mixtures of sugars, has led to the use of high fructose corn syrup as a raw material for the manufacture of pure fructose. The technology was reviewed by M. Seidman.[59]

The role of fructose as a food and in the food processing industry in the United States is not yet clear. Although its desirable physical and chemical characteristics are well documented,[57] its basic attractiveness appears to be related to its relative sweetness when compared with sucrose.[60] In theory, substitution of fructose for sucrose in foods (especially soft drinks) would permit a reduction in the caloric content of a food because of the greater relative sweetening power of fructose. However, experimental studies have demonstrated that fructose sweetness varies with the conditions under which it is used.[61] A theoretical basis explaining this variability in sweetness has been set forth by Shallenberger.[62]

BY-PRODUCTS

Molasses

Molasses, a residual mother liquor from which little or no additional sugars can be recovered economically, is a by-product common to the cane and beet sugar industries, as well as the dextrose industry. Each type of manufacturer has designated this liquid by-product with its own peculiar name: the molasses from cane sugar production is most commonly called blackstrap, that from beets is simply beet molasses, while that from starch hydrolysis is known in the United States as hydrol.

The molasses from each source differs considerably from the others, as shown in Table 9.4.

Blackstrap contains a significant quantity of both sucrose and reducing sugars. A substantial amount of reducing sugar is produced during the processing of cane juice into sugar and molasses. Beet molasses contains primarily sucrose and little or no reducing sugar because of highly alkaline processing conditions that destroy reducing sugars. Hydrol, of course, contains no sucrose. Higher saccharides result from incomplete

TABLE 9.4 Analysis of Molasses from Various Sources

Constituent	*Blackstrap*[63] *(Dry Basis)*	*Beet Molasses*[64] *(Dry Basis)*	*Hydrol*[21] *(Wet Basis)*
Sucrose	37.4	63.0	
Reducing sugars	32.7	0.7	55.0
Higher saccharides		1.5	
Ash	13.5	12.0	7.2
Nitrogen	0.4	2.0	0.07

TABLE 9.5 Cane and Beet Molasses Approximate Production, 1985–89 (Millions of Metric Tons)

Year	*Production*
1985/86	36.9
86/87	39.8
87/88	40.5
88/89	41.7
89/90	43.3

hydrolysis or from the polymerization of "sugar" units during processing.

The world production of blackstrap and beet molasses closely parallels sugar production. Total world production is shown in Table 9.5.

The major outlet for all molasses at the present time is in feed for cattle or other animals. Fermentation into rum, potable ethanol, citric acid, and vinegar continues to be a market for considerable quantities of molasses; and large quantities are used for growing yeasts. However, these uses have diminished in relation to feed use over the years. Beet molasses, because of its high content of nitrogeneous substances, is valued for the production of citric acid by fermentation.

When the price of petroleum is high, conversion of molasses and cane juice into absolute ethanol as partial or total substitution for gasoline in automotive vehicles becomes attractive. Brazil uses more than half of the available cane to produce ethanol, for use mainly as fuel for its cars.[65–67]

Bagasse and Beet Pulp

An important by-product from a cane sugar mill is bagasse, the fibrous portion of cane from which juice is extracted. When it is discharged from the milling train, bagasse contains solid matter (short fibers and the spongy tissue of the pith) and 50 percent by weight of water. Because the average fuel value of ash-free dry bagasse is 8300 Btu/lb, it is primarily used as fuel for the generation of steam in sugar factories. However, when an excess of bagasse is available or alternate fuels are plentiful and reasonably priced, the bagasse is utilized as a raw material for the pulp, paper, paperboard, and wallboard industries, especially in wood-poor areas of the world. Other successful uses of bagasse include its conversion into furfural and its use as a filler for explosives. Today, bagasse is receiving a great amount of attention as a source of sugar and fuel.[68]

Beet pulp is the structural portion of beet roots that remains after spent beets are discharged from a diffuser. As it is highly hydrated at that time, the wet pulp is dewatered in pulp presses. Then it is discharged into wet silos, and it may be sold as wet pulp for animal feed. However, more often it is dried or mixed with molasses and dried, in both of which cases it is used in the manufacture of animal feeds.

Other By-products

When grains such as corn, sorghum, and wheat are used as sources of starch, a number of important by-products are obtained, including steep water, corn oil, gluten, and hulls. Most of these products are used as ingredients in mixed feeds, but corn oil is used widely as a food product. By-products from the other commercial sources of starch are not so important as those from grains, but they usually find their way into feeds if they are salvaged.

TRADE IN SUGAR AND OTHER SWEETENERS

Many countries—especially the tropical, cane-sugar-producing ones—produce more sugar than is consumed locally. A number of beet-sugar-producing countries, particularly in Europe, became self-sufficient in the late 1970s. However, some European countries continue to import raw sugar to be refined. Altogether, 29.3 million metric tons, raw value, of centrifugal sugar were exported globally in 1989–90. Russia, EC-12 countries, the United States, Japan, China, South Korea, and Mexico are the largest importers, with Cuba, Australia, Thailand and Brazil among

the largest exporters of raw sugar. The EC-12 countries are the world's largest exporter of refined white sugar, and the world's second largest sugar exporter, behind Cuba.

Trade in sugar throughout much of the world is influenced by agreements and special trading arrangements. For instance, the United States regulated sugar production, distribution, and importation by means of the Sugar Act of 1948 as amended,[69] but the act was allowed to expire in 1974. From 1981 to the present, the United States has had a price-support program, connected to farm legislation. In 1982, an import quote system went into effect to supplement the price-support program. In 1988, Australia filed a complaint against the United States, charging that the import restriction on sugar violated the General Agreement of Tariffs and Trade (GATT) rules. When the United States was found to be in violation on some of these rules, it implemented a change from an absolute quota to a tariff-rate quota.

The United States also has a program called the General System of Preferences (GSP) that exempts countries (mostly in the Caribbean area) from any duty.

The EC-12 countries import raw sugar within a special arrangement with ACP (Africa, Caribbean, and Pacific) countries.

Since 1986, the world community of nations has tried to expand world trade by formulating new rules and regulations on a wide variety of products within the General Agreement on Tariffs and Trade (GATT). Of all the products discussed, sugar became a major stumbling block, and the end result with sugar was a stalemate. Negotiations were concluded in December 1990.

Other sweeteners, such as dextrose, glucose syrups, honey, maple sugar and syrup, and edible molasses, have been available in the United States for many years. However, only dextrose and glucose syrup have been produced and distributed in substantial quantities in the United States (and many other countries, particularly western Europe) as alternative or supplemental sweeteners.

Per capita consumption of dextrose in the United States has been relatively constant at 4 to 5 lb for many years. The growth in the per capita consumption of starch-derived sweeteners prior to 1970 occurred in glucose syrup (a family of hydrolyzed syrups manufactured from starch), but that growth was limited to less than 2 lb per year. The increased use in the starch-derived sweeteners had no apparent effect on the consumption of cane and beet sugar, which had remained more or less constant at 97 to 99 lb per capita until the mid-1970s when HFCS began to replace sugar significantly in the beverage industry.[39] That change was due largely to the utilization of a new technology (immobilized enzymes), as well as governmental action to protect the domestic sweetener industry. The trend of the changes introduced by this product is shown in Fig. 9.18. HFCS consumption increased from less than one pound per capita in 1968 to more than 48 lb in 1990. For all intents and purposes HFCS is a replacement for invert sugar syrups. HFCS also displaces sucrose syrups in some applications. Those trends are apparent from Fig. 9.18, which shows steady increases in HFCS consumption during the same period. However, the declining consumption of sugar is due to several factors, only some of which are related to the availability of HFCS.

HFCS technology and products have spread to other countries, particularly Canada, Japan, and several countries in western Europe. The pace of commercial development in those other areas does not appear to be so rapid as that in the United States, where a combination of plentiful supplies of starch at reasonable prices and an existing industry based on starch hydrolysis appear to catalyze the development.

ENVIRONMENTAL CONCERNS

The sugar industry, in common with all other sectors of the chemical process industries, must be concerned with protecting the environment. This means that close attention must be paid to all gaseous and liquid effluents, with the aim of minimizing their production and ensuring that they at least meet all minimum standards before being discharged into the environment.

Fig. 9.18. Trends in distribution of sugars in the United States.

It is certain that present standards will not be relaxed, and that they can be expected to become even more restrictive as time goes on. It is therefore essential that older plants and operations be reviewed continually to determine what further steps can be taken to control emissions and reduce their volume. It is assumed that any changes made to existing operations, as well as all new designs, will include careful evaluation of their implications for waste management requirements. Chapters 2 and 3 treat the subjects of waste minimization and wastewater treatment at length and can be useful in planning waste management strategies.

Clarke[72] has written that: "A recent review of the U.S. beet sugar industry emphasizes the increasing severity of impending restrictions. Of particular concern are pesticide and heavy metal residues that can enter groundwater from any source. It is in these areas of water treatment, particularly for refineries and factories in urban areas, that such processes as reverse osmosis treatment may find application."

The following tables from Clarke's paper provide information on the by-products generated in the industry. Some of them are useful and can be sold, whereas others constitute waste streams and must be disposed of in acceptable ways.

Sugarcane Factory By-products

	% on cane
Molasses	2.6%–4% (85 Bx)
Bagasse	23%–35% (50% mois.)
Filter muds	2.5%–4.5% (75% mois.)
Flue gases	100%–130%
Furnace ash	0.3%–0.5%
Factory effluent waters:	0.05–0.1 m^3/ton cane

Byproducts from U.S. Cane Sugar Refineries

Molasses		
Effluent water:		
	Type:	*BOD, mg/l*
	Vessel and filter wash	10–600
	Char wash water	800–2400
	Truck and rail car wash	15,000–18,000
Filter muds	calcium carbonate	
	calcium phosphate	
Spent carbon, bone char, resin		

Clark also refers to several studies related to odor control problems and current and pending environmental controls in the Netherlands and Germany.

REFERENCES

1. Chenu, P. M. A. M., "Alcohol Manufacture in a Sugar Factory," *XVI Congress of International Society of Sugar Cane Technologists*, pp. 3241–3250, Sao Paulo, Brazil, 1978.
2. Tegge, G., "Development and Prospects of Sweeteners Obtained by Glucose Isomerization," *International Sweetener Report No. 1*, H. Ahlfeld (Ed.), pp. 3–8, Dec. 1978.
3. Hugill, J. A. C., "History of the Sugar Industry," in *Sugar: Science and Technology*, G. G. Birch and K. J. Parker, (Eds.), pp. 15–37, Applied Science Publishers Ltd., London, 1979.
4. Meade, G. P., *The Sugar Molecule*, Sugar Research Foundation, N. Y., **11** (1), 11–19 (1959–60).
5. Deere, N., *The History of Sugar*, Vol. I, Chapman and hall, London, 1949.
6. Hugot, E., and Jenkins, G. H., *Handbook of Cane Sugar Engineering*, American Elsevier, New York, 1960.
7. McGinnis, R. A., *Beet Sugar Technology*, Van Nostrand Reinhold, New York, 1951.
8. Clarke, M. E., "Cane Sugar Process and By-product Streams," Iowa Biotechnology Consortium, Mar. 1991.
9. Deere, N., *Cane Sugar*, Norman Rodger, London, 1921.
10. Ducasse, V., "Heavy Duty Chopper Fiberiser," *XVI Congress of International Society of Sugar Cane Technologists*, pp. 1568–1601, Sao Paulo, Brazil, 1978.
11. Rodriquez, L., Garcia, E. G., and Alba, A. C., "New Concept in Cane Preparation," *Sugar y Azucar*, pp. 55–59 (Nov. 1977).
12. Battelle Staff, "Sugar Crops as a Source of Fuels," Vol. II, Report to U.S. Department of Energy by Battelle Laboratories, pp. 51–65, Aug. 1978.
13. Meade, G. P., *Cane Sugar Handbook*, John Wiley & Sons, New York, 1964.
14. "BMA Cane Diffusion," *Sugar y Azucar*, p. 22 (July 1970).
15. Whayman, E., and Crees, O. L., "Mechanistic Studies of Cane Mud Flocculation," *XVI Congress of International Society of Sugar Cane Technologists*, pp. 1175–1182, Sao Paulo, Brazil, 1978.
16. Baikow, V., *Manufacture and Refining of Raw Cane Sugar*. American Elsevier, New York, 1967.
17. Honig, P., *Principles of Sugar Technology*, Vol. II, American Elsevier, New York, 1959.
18. Rein, P. N., "A Review of Experience with Continuous Pans in Tongatt-Hulett Sugar," *Proc.* S. African Sugar Technologists' Assn., 1986.
19. Clarke, M. E., Private communication, 1991.
20. Silver, B., "Continuous Vacuum System," *Sugar y Azucar*, p. 36 (1970).
21. Honig, P., *Principles of Sugar Technology*, Vol. III, American Elsevier, New York, 1963.
22. Clarke, M. E., "Recent Developments in Technology of Cane Sugar Production," *Sugar y Azucar*, pp. 6–7 (1991).
23. Honig, P., *Principles of Sugar Technology*, Vol. I, American Elsevier, New York, 1953.
24. Rundell, J. T., Rich, C. V., and Norcott, A. E., "A New Automatic Defecation–Decolourization Station Using the Talofloc Process," *Pub. of Technical Papers and Proceedings of Sugar Industry Technologists*, **32**, 58–70 (1973).
25. Pollio, F. X., and McGarvey, F. X., "Use of Ion Exchange in Cane Sugar Processing," *XVI Congress of International Society of Sugar Cane Technologists*, pp. 2729–2741, Sao Paulo, Brazil, 1978.
26. Dmitrovsky, M., "Sucrose Process for Evaporative Continuous Crystallization," *Technical Papers and Proceedings of the Sugar Industry Technologists, Inc.*, Vol. 38, pp. 291–302, 1979.
27. Rodgers, T., and Lewis, C., "Drying of Sugar and Its Effect on Bulk Handling," 15th Tech. Conf. Report, British Sugar Corp., *International Sugar Journal*, **64**, 359–362 (1962).
28. Clarke, M. E., "Recent Developments in Cane Sugar Production," *Sugar y Azucar*, p. 7 (1991).
29. Bennett, M. C., Elvin, J. R., Heineman, H. W. B., Pottage, P. R., and Rundell, J. T., "Syrup Clarification for Improving Sugar Quality and Yield," *XVI Congress of International Society of Sugar Cane Technologists*, pp. 2797–2809, Sao Paulo, Brazil, 1978.
30. Leao, F. M. D. "Production of Amorphous Refined Sugar in Brazil," *XVI Congress of International Society of Sugar Cane Technologists*, pp. 2797–2809, Sao Paulo, Brazil, 1978.
31. Bickert, W. G., Bakker-Arkema, F. W., and Dexter, S. T., "Refrigerated Air Cooling of Sugar Beets," *Journal of the American Society of Beet Sugar Technologists*, **14** (7), 547–554.
32. Kearney, M. "Simulated Moving-Bed Technology Applied to Chromatographic Recovery of Sucrose from Sucrose Syrups," *Proc.* 1990 Conference on Sugar Processing Research, pp. 291–384, 1990.
33. Schneider, H. G., and Mikule, J., "Recovery of Sugar from Beet Molasses by the P & L Exclusion Process," *International Sugar Journal*, **77**, 259–264, 294–298 (1975).
34. Munir, M., "Molasses Sugar Recovery by Liquid Distribution Chromatography," *International Sugar Journal*, **78**, 100–106 (1976).
35. Hongisto, H., and Hikkala, H., "Desugarization of Cane Molasses by the Finnsugar Chromatographic Separation Process," *XVI Congress of International Society of Sugar Cane Technologists*, pp. 3031–3037, Sao Paulo, Brazil, 1978.

36. Rapaport, R. M., Monti, A., Moroz, R. D., and Broeg, C. B., "Process for Recovering Useful Products from Carbohydrate-Containing Materials," U.S. Patent No. 4,101,338, July 18, 1978.
37. "Thick Juice for Off-season Processing," *Beet Sugar*, **57**, 46–49 (Jan. 1962).
38. Licht, F. O., *International Sugar Reports*, **102** (20), 1–4 (1970).
39. U.S. Department of Agriculture, *Sugar and Sweetener Report*, SSR, Vol. 5, No. 2, p. 33, Washington, DC, Feb. 1980.
40. Hickson, J. L., "The Potential for Industrial Uses of Sucrose," in *Sugar: Science and Technology*, G. G. Birch and K. J. Parker (Eds.), pp. 151–180, Applied Science Publishers Ltd., London, 1979.
41. Kollonitsch, V., *Sucrose Chemicals*, pp. 17–105, The International Sugar Research Foundation, Inc., 1970.
42. Kollonitsch, V., *Sucrose Chemicals*, pp. 173–221, The International Sugar Research Foundation, Inc., 1970.
43. Vlitos, A. J., World Sugar Research Organization Symposium, Caracas, Venezuela, *Sucrochemistry—New Products from Sugar*, published by World Sugar Research Organization, London.
44. Khan, R., "Advances in Sucrose Chemistry," in *Sugar: Science and Technology*, G. G. Birch and K. J. Parker (Eds.), pp. 181–210, Applied Science Publishers Ltd., London, 1979.
45. Kooreman, J. A., "Physical and Chemical Characteristics of Enzyme Converted Syrup," *Manufacturing Confectioner*, pp. 35–90 (June 1955).
46. MacAllister, R. V., *Advances in Carbohydrate Chemistry and Biochemistry*, Vol. 36, R. S. Tipson and D. Horton (Eds.), Academic Press, New York, 1979.
47. Reed, G., *Enzymes in Food Processing*, Academic Press, New York, 1966.
48. Kerr, R. W., *Chemistry and Industry of Starch*, Academic Press, New York, 1944.
49. Dean, G. R., and Gottfried, J. B., *Advances in Carbohydrate Chemistry*, Vol. 5, Academic Press, New York, 1950.
50. Vabo, B., "Enzyme Systems for Production of HFCS Syrups," presented at International Biochemical Symposium, Stone and Webster Engineering Corp., Chicago, IL, Oct. 1976.
51. Hupkes, J. V., "Practical Conditions for the Use of Gist-Brocades' Immobilized Glucose Isomerase: Maxazyme (R) GI-IMMOB," presented at International Biochemical Symposium, Toronto, Oct. 1977.
52. Newton, J. M., and Wardrip, E. K., "High Fructose Corn Syrup," *Symposium: Sweeteners*, G. E. Inglett (Ed.), The Avi Publishing Co., Inc., Westport, CT, 1975.
53. Nordlund, D. E., "High Fructose Syrups: The Competition to Sugar," *XVII Congress of International Society of Sugar Cane Technologists*, Manila, 1980.
54. Hoare, M., "Implications of Immobilized Glucose Isomerase Specifications for the Economics of High Fructose Glucose Syrup Production," *International Sweetener Report No. 1*, H. Ahlfeld (Ed.), pp. 23–32, Dec. 1978.
55. Keller, H. W., and Reents, A. C., "Production of Very Enriched Fructose Corn Syrup (VEFCS) with the Technisweet (TM) Systems," presented at International Biochemical Symposium, Toronto, 1977.
56. Broughton, D. B., Beiser, H. J., Berg, R. C., and Connell, E. D., "High Purity Fructose via Continuous Adsorptive Separation," *La Sucrerie Belge*, **96**, pp. 155–162 (May 1977).
57. Doty, T. E., and Vanninen, E., "Crystalline Fructose: Use as a Food Ingredient Expected to Increase," *Food Technology*, pp. 34–38 (Nov. 1975).
58. Bates, F. J., and Associates, *Polarimetry and Saccharimetry of Sugars*, Circular C440, pp. 398–403, Nat. Bur. of Standards, U.S. Dept. of Agriculture, Washington, DC, 1942.
59. Seidman, M., "New Technological Development in D-Fructose Production," in *Developments in Food Carbohydrate-1*, G. G. Birch, and R. S. Shallenberger (Eds.), pp. 19–42, Applied Science Publishers, London, 1977.
60. Harris, W. E., Rubino, J. A., and McNutt, J. W., *Food Processing Industry*, pp. 28–29 (Feb. 1978).
61. Cardello, A. V., Hunt, D., and Mann, B., *Journal of Food Science*, **44**, 748–751 (1979).
62. Shallenberger, R. S., "Predicting Sweetness from Chemical Structure and Knowledge of Chemoreception," *Food Technology*, pp. 65–66 (Jan. 1980).
63. Binkley, W. W., and Wolfram, M. L., "Composition of Cane Juice and Final Molasses," *Advances in Carbohydrate Chemistry*, Vol. 8, Academic Press, New York, 1953.
64. Stark, J. B., and McCready, R. M., "The Relation of Beet Molasses Composition to True Purity," *Journal of American Society of Beet Sugar Technologists*, **15**, 61–72 (1968).
65. Lipinsky, E. S., Nathan, R. A., Sheppard, W. J., McClure, T. A., Lawton, W. T., and Itis, J. L., "Systems Study of Fuels from Sugarcane, Sweet Sorghum and Sugar Beets," Vol. 1, Comprehensive Evaluation, Report to U.S. Energy Research and Development Administration by Columbus Laboratories, Columbus, OH (Mar. 15, 1977).
66. Kampen, W. H., "Ethyl Alcohol—the Automobile Fuel of the Future," *Sugar y Azucar*, pp. 18–30 (Apr. 1978).
67. Hopkinson, C. S., Jr. and Day, D. W., Jr. "Net Energy of Alcohol Production from Sugarcane," *Science*, **207**, 302–303 (Jan. 1980).
68. Tsao, G. T., "Sugar and Sweeteners from Sugarcane," *International Sweetener Report No. 2*, H. Ahlfeld (Ed.), F. O. Licht GmbH, Ratzeburg, Fed. Repub. of Germany, Sept. 1979.
69. U.S. Congress, Public Law 339.
70. Corn Industries Research Foundation, *Corn Syrups and Sugars*. Washington, DC, 1965.
71. Clarke, M. E., "Sugar Industry Products and Processes," *Sugar y Azucar Yearbook*, p. 27, (1989).

72. Clarke, M. E., "Sugar Industry Products and Processes," *Sugar y Azucar Yearbook*, p. 30, (1989).
73. Hagelbert, G. B., and Ahlfeld, H., "The Sugar Industry in a Changing Economic Climate," *XVII Congress of International Society of Sugar Cane Technologists*, Manila, 1980.
74. U.S. Department of Agriculture, *Molasses Market News Annual Summary*. Washington, DC, 1969.
75. U.S. Department of Agriculture, *Sugar and Sweetener Situation and Outlook Yearbook*, SSR, Vol. 2, pp. 33–34, Washington, DC, June 1991.
76. U.S. Department of Agriculture, *Sugar and Sweetener Report*, SSR, Vol. 4, No. 5, Washington, DC, May 1979.

10

Phosphorus and Phosphates

G. A. Gruber* and D. W. Leyshon*

INTRODUCTION

Phosphates, compounds of the element phosphorus, are produced from relatively abundant supplies of phosphate rock, containing apatite, $Ca_3(PO_4) \cdot CaF_2$, together with up to 40 percent of other minerals and impurities.

The major use of phosphate, referred to generally as P_2O_5, its anhydride, is as one of the three major plant foods, N, P, and K. The technology of the extraction of phosphate rock from its ore and its subsequent processing to fertilizer materials and industrial chemicals is a relatively mature art. Phosphate in the form of single superphosphate, a mixture of monocalcium monohydrate and gypsum, formed by the reaction of sulfuric acid with phosphate rock, has been used as a fertilizer since the mid-1800s. Phosphoric acid, derived by the treatment of phosphate rock with sulfuric acid so as to produce gypsum in a separable form, was produced in many locations by batch and countercurrent decantation methods in the 1920s.

This latter process is known as the "wet" process to distinguish it from the thermal or furnace process, in which elemental phosphorus is volatized and subsequently burned to P_2O_5 and hydrolyzed to H_3PO_4.

The "wet" processes, used for the bulk of today's phosphoric acid and fertilizer production, trace their origin to the early 1930s.[1]

This text has utilized portions of the chapter on phosphorus and phosphates from the eighth edition by J. H. McLellan, particularly the description of the furnace production of phosphorus and its derivative phosphates, which has remained relatively unchanged over the last decade.

Furnace processing, in contrast to the "wet" process, has been the method most often used for the production of higher-purity industrial and food-grade phosphates. Recently, however, solvent extraction technology has advanced to the level that food grade phosphates are being produced from relatively impure wet process phosphoric acid.

PHOSPHATE ROCK

Naturally occurring phosphate products having commercial value are classified as phosphate rock. The grade or phosphate content of these

*Jacobs Engineering Group, Inc..

TABLE 10.1 Phosphate Rock Value (United States Marketable Production)

Year	*Total Value billion $*	*Unit Value $ per metric ton*
1980	1.148	21.34
1981	1.368	24.49
1982	1.204	26.44
1983	0.918	24.88
1984	1.182	24.03
1985	1.236	24.31
1986	0.897	22.25
1987	0.793	19.37
1988	0.888	19.56
1989	1.084	21.76

Source: Mineral Industry Surveys, U.S. Department of Interior, Bureau of Mines.

products traditionally has been reported as $Ca_3(PO_4)_2$, which may be called bone phosphate of lime (BPL), tricalcium phosphate (TCP), or triphosphate of lime (TPL). Factors equating the traditional rock analysis to the analytical terms commonly used in chemical and fertilizer technologies are: 100.00 percent BPL = 45.77 percent P_2O_5 = 19.96 percent P.

International trade of phosphate rock is based on dry metric tons, whereas U.S. domestic sales are on a short ton basis for either dry rock or wet rock. Chemical impurities such as CaO, MgO, Al_2O_3, Fe_2O_3, Na_2O, K_2O, F, Cl, and organic matter may be of importance, depending on the intended use of the phosphate rock.

Value

The value of phosphate rock, like that of many industrial minerals, is influenced by supply and demand, product specifications, and freight costs. Values of U.S. marketable production of phosphate rock from selected years are listed in Table 10.1.

Production

Phosphorus, in addition to nitrogen and potassium, is an essential ingredient of modern high analysis fertilizers. An ever-increasing global demand for food has resulted in increased fertilizer use and increased phosphate rock consumption. The data presented in Table 10.2 show a marked increase in world phosphate production, equivalent to an annual growth rate of 3.5 percent from 1970 to 1989.

Many countries use indigenous phsophate rock as a source of phosphorus and phosphates for industrial chemicals and fertilizers. Few countries, however, are self-sufficient, and supplemental sources of phosphate rock are essential.

Comparative statistics for the six leading exporters of phosphate rock are given in Table 10.3.

Resources

McKelvey[2] reported that the earth's crust contains an average of about 0.27 percent P_2O_5, most of which occurs as species of

TABLE 10.2 World Production of Phosphate Rock Production (Millions of Short Tons)

	1970	*1975*	*1980*	*1985*	*1989*
United States	38.7	48.8	58.8	54.6	54.2
Soviet Union	22.5	26.6	27.2	37.8	42.9
Morocco	12.6	14.9	20.8	22.9	19.8
China	1.3	3.7	11.8	7.7	18.7
Tunisia	3.3	3.8	5.1	5.0	7.3
Jordan	1.3	1.5	4.7	6.7	7.3
Other	14.1	19.3	14.4	28.7	29.6
Total World	93.8	118.6	142.8	163.4	179.8

Source: "Central Florida Phosphate Industry 1990 Status and Outlook," Zellars-Williams Company.

TABLE 10.3 Major Exporters of Phosphate Rock (Million Short Tons of Exported Phosphate Rock)

	1988	*1989*
Morocco	15.7	13.6
United States	10.2	9.2
Jordan	6.2	6.5
Soviet Union	3.5	3.5
Togo	3.2	3.7
Senegal	2.0	1.6

Source: "Central Florida Phosphate Industry 1990 Status and Outlook," Zellars-Williams Company.

apatite. Igneous rocks make up about 95 percent of the earth's crust but contain few apatite deposits of commercial value. Sedimentary rocks, which make up about 75 percent of the earth's surface, host the majority of commercially exploited phosphate deposits. Guano deposits formed from sea bird droppings and from bat droppings are of minor importance, as are guano-related deposits.

Naturally occurring phosphates exist as, or orginated as, accessory minerals in igneous rocks. Prolonged weathering gradually decomposes apatite so that its phosphate content becomes soluble. Eventually the soluble phosphate concentrates in the world's oceans where it is precipitated through chemical or biochemical reactions to form sedimentary marine phosphorites. Precambrian sedimentary phosphate deposits of commercial value exist in China; however, marine phosphorite deposition was more common in the Mesozoic and Cenozoic eras. deBelinko[3] has used paleogeographic reconstruction to summarize major periods of phosphate deposition and formation. The world's resources of phosphatic material amount to hundreds of billions of tons. According to Bernardi and Hall,[4] secondary enrichment through weathering, water action, and phosphatization is an important aspect in the formation of deposits.

Phosphatic resources normally are quantified in terms of in situ tons of phosphate material, without regard to economic and chemical criteria. Northolt, Sheldon, and Davidson[5] state that the identified phosphate resources in North America total over 35 billion metric tons. Phosphate ore reserves should be quantified in terms of phosphate rock recoverable at specified chemical and economic criteria. The definition of phosphate ore reserves normally requires an integrated program of geological exploration, laboratory testing, and application of criteria.

PHOSPHATE ORES

Phosphate ores occur in igneous deposits, sedimentary deposits, and guano-related deposits. According to Cathcart,[6] the percentages of phosphate rock produced from igneous, sedimentary, and guano-related deposits are 20, 75, and 5, respectively.

The types of igneous deposits in which commercially important concentrations of apatite have been found are nepheline-seyenite and carbonatites. The world's largest apatite deposit, located in Russia's Kola Peninsula, is associated with nepheline-seyenite. Carbonatite deposits that are mined for their phosphate content include Siilinjarvi in Finland, Araxa and Jacupiranga in Brazil, and Phalaborwa in South Africa. By-product recovery from carbonatite ores is common. At Siilinjarvi, by-product calcite is recovered for use as agricultural lime, whereas copper concentrate and baddeleyite (zirconium oxide) are by-products at Phalaborwa.

Sedimentary phosphate ores may be distinguished by their gangue minerals. Typically, phosphorite ores are characterized as having siliceous gangue or carbonate gangue. Soluble salts and organic material are also of concern. The Bone Valley formation in central Florida, which has been mined for one hundred years, contains phosphate ore with siliceous gangue comprised of quartz and clay minerals. The Hawthorn formation, also in central Florida, contains both siliceous gangue and carbonate gangue.

Examples of phosphorites containing carbonate gangue are Djebel Onk in Algeria and Akashat in Iraq. The usual carbonate minerals are calcite and dolomite. By-products from

phosphorite ores are not common; however, uranium (U_3O_8) has been extracted commercially from phosphoric acid.

The common phosphate minerals are of the apatite group, with fluorapatite, hydroxyapatite, and chlorapatite in igneous rocks, and carbonate fluorapatite or francolite in sedimentary rocks. McClellan,[7] has quantified the chemical composition of the end members of the fluorapatite-francolite series as shown in Table 10.4. The empirical formulas for the end members are:

$$\text{Fluoroapatite} = Ca_{10}(PO_4)_6F_2$$

$$\text{Francolite} = Ca_{10-a-b}Na_aMg(PO_4)_{6-x}(CO_3)_xF_{0.4x}F_2$$

TABLE 10.4 Mineral Composition of the End-Members of the Fluorapatite–Francolite Series

Constituent	*Fluorapatite*	*Francolite*
% CaO	55.6	55.1
% P_2O_5	42.2	34.0
% CO_2	0	6.3
% F	3.77	5.04
% Na_2O	0	1.4
% MgO	0	0.7

Source: McClellan.[7]

TABLE 10.5 Typical Phosphate Rock Analysis, % Dry Basis[10]

	Central Florida 68 BPL	*Morocco K-20, 70–71 BPL*
H_2O	(1.3)	(1.3)
P_2O_5	31.2	32.0
SO_3	1.1	1.88
F	3.7	3.9
SiO_3	9.0	2.1
CO_2	3.8	6.0
Cl	0.01	0.02
CaO	46.7	51.2
Al_2O_3	1.1	0.3
Fe_2O_3	1.5	0.2
MgO	0.5	0.5
Na_2O	0.5	0.9
K_2O	0.11	0.07
C(org.)	0.28	0.22
Cd	5	18
U_3O_8	185	134

As the carbonate content of the ore increases, the substitution of carbonate in the francolite tends to increase. It is common for trace quantities of heavy metals, toxic elements such as cadmium, and radionuclides such as uranium to exist in phosphate rock owing to substitution in the crystal lattice (see Table 10.5). Usually the quantities of these elements are insufficient to be of concern.

Mining

Phosphate ores are extracted by underground methods from deep deposits and by surface mining methods from shallow deposits—the former being more costly and thus less common than the latter.

Tunisia has seven underground mines; six use room and pillar mining techniques, while the seventh extracts the ore by longwall mining. Room and pillar mining also is utilized at Warm Springs in Montana, Khouribga in Morocco, and San Juan de la Costa, Baja California. Shrinkage stopping is the mining method at Kola, Russia.

Many shallow phosphate deposits have unconsolidated ore and are covered by unconsolidated overburden. Large electric walking draglines are ideally suited for such deposits, as evidenced by their use at large-capacity phosphate mines in the United States, Morocco, Jordan, Senegal, and Togo. Some sedimentary deposits have overburden and ore ranging from partially to fully consolidated. For these deposits, ripping or drilling and blasting are required to fragment the overburden and ore so that they can be excavated. In addition to draglines, power shovels, backhoes, and wheeled loaders commonly are used for excavation.

Bucketwheel excavators and cutter head dredges have not yet found wide acceptance in phosphate mining. The Lee Creek phosphate mine in North Carolina has used both cutter head dredges and bucketwheel excavators to remove the upper portion of the overburden, with the remaining overburden and ore excavated by dragline. In Togo, a combination of bucketwheel excavator and dragline is used for stripping overburden and mining ore. The

Wingate Creek phosphate mine in central Florida uses two cutter head dredges; one for stripping overburden and the second for mining ore and pumping the slurried ore to the beneficiation plant.

Transportation of phosphate ore from the mine to the beneficiation plant is an important component of production. Typical modes of transportation are slurry pipeline, truck, railroad, and belt conveyor. Slurry pipelines 18 to 22 inches in diameter operating at less than 100 psi are used exclusively in central Florida, North Florida, and North Carolina. Conventional centrifugal slurry pumps are spaced at approximately 1000-yard intervals to extend the pipelines up to 10 miles or more. Phosphate mines in Tennessee use trucks to haul ore several miles from scattered mine areas to a centralized beneficiation plant. In Iraq, phosphate ore is shipped from the Akashat mine 150 km by railroad to the plant at Al Qaim. Phosphate ore from the Bou Craa mine in Morocco (formerly the Spanish Sahara) is transported 30 km by belt conveyor to the plant at La Youn.

Beneficiation

A vast majority of phosphate ores yield commercial phosphate rock only after the removal of impurities by beneficiation. Also known as mineral dressing and ore processing, beneficiation may involve a variety of operations such as size reduction, size separation, mineral separation, and thermal processing.

The usual first beneficiation operation is size reduction by crushing, grinding, or disaggregation from washing and scrubbing. This step liberates the mineral species so that they later can be separated. Size separation usually follows size reduction. Frequently carbonate and silicate gangue is more indurated than the phosphate mineral and can be rejected as oversize by wet or dry screening. Argillaceous material and some microcrystalline carbonates are smaller than the phosphate mineral grains and can be rejected as fines by either wet or dry classification. A few phosphate ores contain soluble chlorides, and some ores are beneficiated with seawater. The phosphate rock recovered from these ores must be washed or rinsed with fresh water to remove chlorides.

When the gangue minerals have the same particle size as the phosphate grains, process techniques such as froth flotation, magnetic separation, and heavy media separation are required to reject the impurities. Froth flotation is the proven method for rejecting silica and silicate gangue from sedimentary phosphate ore. A two-stage flotation process consisting of anionic flotation to recover a low-grade phosphate concentrate with fatty acid collectors and fuel oil, followed by cationic flotation with amine collectors to reject quartz and obtain a high-grade phosphate concentrate, is used in Florida and North Carolina. A typical flow diagram for beneficiating Florida phosphate ore is given in Fig. 10.1. Several Florida phosphate flotation plants installed column flotation cells during the period 1978 to 1984. These cells have met with varying degrees of success, but remain strong contenders for plant modifications or expansions.

Igneous phosphate ores having carbonate gangue also are successfully treated by froth flotation at Siilinjarvi, Jacupiranga, and Phalaborwa. Froth flotation has not yet been widely adopted for sedimentary phosphate ores containing carbonate gangue. Commercial plants where flotation systems have been installed to separate carbonates from phosphates are Kara Tau in Russia, Wangji in China, and Conda in Idaho. Houot gave a good review of phosphate flotation technology.[8]

Magnetite, an accessory mineral in some of the igneous ore deposits, is easily removed from phosphate rock by low-intensity magnetic separators. In Brazil, magnetic separation is required at Arafertil, Camig, and Fosfago. Heavy media separation has been proposed as a method of rejecting low-density carbonate gangue from phosphate pebble recovered from the southern extension of the central Florida district.[9]

Thermal drying to reduce the moisture to about 3 percent, or less, frequently is necessary to prepare phosphate rock for export sales. Rotary dryers, fluid bed dryers, and, to a lesser

Fig. 10.1. Typical flow diagram, Florida phosphate rock concentration. (*Courtesy Jacobs Engineering Group, Inc.*)

extent, flash dryers are used for this purpose. Organic matter may be removed from phosphate rock by calcining at temperatures of about 1300 to 1500°F. Organic removal by calcination is practiced at Lee Creek, North Carolina and at Youssoufia, Morocco. Calcination of phosphate rock at about 1800°F decomposes calcium and magnesium carbonates and drives off carbon dioxide. The calcium and magnesium oxides may be removed by quenching and slaking of the calcine followed by size separation to reject the hydroxides. The removal of carbonates by calcination with subsequent quenching and slaking is accomplished at Al Qaim, Iraq and Djebel Onk, Algeria. Phosphate rock may be defluorinated by calcining at about 2400°F.

ELEMENTAL PHOSPHORUS AND PHOSPHORIC ACID

Phosphate rock is converted into usable chemicals by two methods. In one, the rock is charged to an electric furnace with silica and coke to produce elemental phosphorus. The phosphorus then is converted into phosphoric acid and other compounds. In the other, the phosphate rock is reacted with sulfuric acid in a medium of phosphoric acid and gypsum crystals to form dilute, impure phosphoric acid. The acid is separated and used to make fertilizers. This is known as the wet process method.

The furnace or thermal process is shown in Fig. 10.2. The approximate reaction is:

$$2Ca_3(PO_4)_2 + 6SiO_2 + 10C \rightarrow 6CaSiO_3 + P_4 + 10CO$$

The charge or furnace burden must have certain physical and chemical properties. The particles should be dry, durable masses in the 1/4 to 3/4 inch size range that will not break up or segregate in the furnace. Phosphatic shale is siliceous and is a suitable material, but usually it must be sintered or briquetted before use. Quartz gravel is a good source of supplementary silica, and metallurgical coke is the form of carbon used.

The energy required is about 12,000 kWh per ton of phosphorus produced. The most modern furnaces are quite large, with ratings in the range of 60,000 kVA.

Fig. 10.2. Electric furnace process for production of elemental phosphorus. (*Courtesy Tennessee Valley Authority.*)

The furnace is a water-cooled circular steel shell lined with refractories. The bottom or hearth is made of thick blocks of carbon. Three large carbon electrodes are mounted through the roof of the furnace and are slowly lowered into the charge as the bottom ends are consumed. A three-phase electrical system is used, and the electrodes must be precisely positioned to balance the phases and maintain a high power factor.

The phosphorus leaves the furnace as a vapor. After passing through an electrostatic precipitator, it is condensed by direct contact with water. Then the liquid phosphorus is filtered or centrifuged to remove dust particles and is shipped in tank cars. CO discharged from the furnace is used as fuel in the feed preparation operations or is burned in a flare.

In 1990, there were about eight plants in operation, some with multiple furnaces. No new electric furnaces have been built in the United States since 1968; moreover, many units have been closed down in recent years. New emission standards, the high capital and operating costs, and competitive wet acid purification technology make it unlikely that new furnaces will be built in the near future. See the discussion of purified phosphoric acid later in this chapter.

Furnace Phosphoric Acid

Phosphoric acid of high purity is made by burning phosphorus with air and hydrating the resulting P_2O_5 with water, according to the reaction:

$$P_4 + 5O_2 + 6H_2O \rightarrow 4H_3PO_4$$

Equipment for this operation has been constructed from both carbon blocks and stainless steel. The water quantity is controlled to give product acid corresponding to 75 or 85 percent H_3PO_4. If food-grade acid is specified, some additional chemical treatment of the acid may be required. For example, slight traces of arsenic are removed by treatment with H_2S.

If even less water is used for hydration, a product known as polyphosphoric acid results. Ordinary phosphoric acid is a solution of the

monomer, H_3PO_4, in water, and is called orthophosphoric acid. If a molecule of water is removed between two orthophosphoric molecules, the dimer, pyrophosphoric acid, $H_4P_2O_7$, is formed. Similarly, the trimer and higher polymers can be made. Superphosphoric acid is a mixture of orthophosphoric acid and polyphosphoric acid. Thermal phosphoric acid has become too expensive for use in the fertilizer industry.

Industrial Phosphates

A portion of the phosphorus currently produced in the United States is consumed to make compounds such as phosphorus pentoxide, phosphoric trichloride, and phosphorus pentasulfide, which find use for the preparation of drying agents, plasticizers, oil additives, fire retardants, and insecticides. Most of the phosphorus, however, is converted into orthophosphoric acid, some of which is used in soft drinks, candy, baked goods, and various other food products. Furnace-grade acid finds wide use in metal-treating methods.

The cheapest and most important salts of phosphoric acid are the sodium salts, made by reacting the acid with sodium carbonate or sodium hydroxide. Sodium phosphates may be classified in a general way as (1) orthophosphates, (2) crystalline condensed phosphates, and (3) glassy condensed phosphates.

Three sodium orthophosphates can be prepared, depending on whether one, two, or three hydrogen atoms are replaced by sodium. Monosodium phosphate is formed in the following reaction:

$$2H_3PO_4 + Na_2CO_3 \rightarrow 2NaH_2PO_4 + H_2O + CO_2$$

Sodium carbonate also is used to make disodium phosphate, Na_2HPO_4, but sodium hydroxide must be used to replace the third hydrogen in trisodium phosphate, Na_3PO_4. The orthophosphates have a wide range of uses in industry. Sodium phosphates are made abroad in large quantities from wet process phosphoric acid by stepwise neutralization. Impurities must be purged from the solution and the salts purified by recrystallization.

TABLE 10.6 Estimated World Consumption of STPP in Soaps, Detergents, and Cleaners 1974–87 (Thousand Tonnes Product)

	1974	*1980*	*1985*	*1987*
West Europe	1000	960	930	725
East Europe	270	340	380	395
North America	825	680	580	555
Africa	50	95	120	130
Latin America	210	275	320	350
Asia and Oceania	210	370	425	460
World	2565	2720	2755	2615

Courtesy British Sulphur Corporation.

Condensed phosphates are made by eliminating water from orthophosphates. The most important compound is sodium tripolyphosphate (STPP), made according to the following reaction:

$$2Na_2HPO_4 + NaH_2PO_4 \xrightarrow{-H_2O} Na_5P_3O_{10}$$

The most important use of sodium tripolyphosphate is as a builder of detergents. However, legislative restrictions on the use of phosphorus compounds in household detergents have caused a worldwide flattening of consumption. See Table 10.6.

Glassy condensed phosphates are represented by sodium hexametaphosphate, in which the O:P ratio is 3:1. There can be considerable variation in the Na_2O to P_2O_5 ratio.

The principal use of the condensed phosphates is to sequester metallic ions in water. They form water-soluble complexes with the metals and prevent metallic compounds from precipitating to cause discoloration, scale, and sludges.

Wet Process Phosphoric Acid

In the wet process, phosphate rock is reacted in a slurry of phosphoric acid and gypsum containing a controlled quantity of sulfuric acid. The simplified reaction is as follows:

$$Ca_3(PO_4)_2 + 3H_2SO_4 \rightarrow CaSO_4 + 2H_3PO_4$$

Until 1969, virtually all wet process acid was made at strengths of 26 to 32 percent

P_2O_5, separating the calcium sulfate as gypsum containing two molecules of water.

Since 1969, a substantial number of commercial hemihydrate process plants have been built in various modes. In the hemihydrate process, reaction conditions are higher in temperature and phosphoric acid strength so that the stable solid phase is calcium sulfate with one half molecule of water.

The hemihydrate process may be a single-stage process in which the hemihydrate solids are the waste product, or the process can have a second step in which the hemihydrate is recrystallized to gypsum.

In the hemihydrate process, acid strengths of 38 percent P_2O_5 to about 44 percent P_2O_5 normally are produced although strengths up to 50 percent can be produced under somewhat more difficult circumstances requiring more filter area.

The neutralization reaction of the above equation is conducted in one or more strongly agitated reaction vessels, whether in a gypsum or in a hemihydrate mode. The system is highly exothermic and the slurry is cooled. Slurry temperatures are maintained at 80 to 85°C for dihydrate processing, 95 to 100°C for hemihydrate.

During the reaction of phosphate rock with sulfuric acid, fluorine is evolved and must be scrubbed from the vent gas.

Table 10.7 shows production of phosphoric acid in the recent past.

Dihydrate Process

The conventional dihydrate process, as operated over the last 60 years, was first demonstrated by the Dorr Company in the Cominco plant at Trail, British Columbia in 1931. The principles discovered at that time for making an acid strength of up to 32 percent P_2O_5 and a highly filterable gypsum crystal are still used. The salient features of the process included maintaining H_2SO_4 content in the digestion acid of about 2.0 to 3.0 percent. Reaction slurry was recycled at a rate of 10 to 15 to one relative to the volume of product slurry sent to the filter. The relatively higher recirculation of seed gypsum than had previously been practiced, together with control of the free sulfate, resulted in product acid well above the previous 22 percent P_2O_5 possible with batch and countercurrent decantation systems.

If the sulfate in the solution is allowed to exceed certain limits, depending on rock reactivity and particle size, the rock becomes coated with gypsum, and the reaction becomes uncontrollable. If the sulfate level is too low, the precipitated gypsum filters poorly and contains excessive quantitites of crystallized (solid solution) P_2O_5, leading to excessive P_2O_5 losses.

Over the years, the process has evolved to employ increased slurry recirculation and different reactor configurations. In addition, several different filter designs are available.

TABLE 10.7 Annual Production of Phosphoric Acid, TPY P_2O_5

	1980	*1985*	*1988*
West Europe	3,669,700	3,442,863	3,322,735
East Europe	3,466,000	4,850,800	6,305,597
Africa	2,043,700	2,762,629	4,661,621
N. Amercia	10,016,800	9,600,200	10,411,968
Mexico	260,400	310,000	474,000
S. America	303,100	695,375	762,806
Middle East	157,000	712,005	813,354
China	9,000	31,500	50,000
South and East Asia	1,414,300	1,637,542	1,637,380
Australia	180,000	133,438	107,650
	21,520,000	24,176,352	28,547,111

Courtesy British Sulphur Corporation.

TABLE 10.8 Phosphoric Acid Plants, Worldwide Dihydrate Process, 1990

Process	*Number of Lines 450 MTPD or Larger*
Prayon	25–30
Rhone Poulenc	15–20
Nissan H	9
Jacobs/Dorr-Oliver	8
Badger-Isothermal	4

The principal processes currently in use are shown in Table 10.8. Over 90 percent of the world's phosphoric acid is produced by the dihydrate method, but there is likely to be increased production by the hemihydrate method because of advantages in some situations. The dihydrate processes listed have been modified over the years so that many slightly different configurations of the same process may exist.

The unique feature of each process is the reactor system configuration. The reactor or reactors normally provide from 3 to 6 hours' detention for the gypsum slurry. This is about 1.0 to 2.0 m^3 of reactor volume per ton of P_2O_5/day, meaning relatively large vessels totaling 1500 m^3 to over 2000 m^3 in size for large plants. The individual reactor systems are described briefly below.

The objective of the reaction system is to produce a highly filterable gypsum crystal that washes well, and that also contains a minimum amount of insoluble P_2O_5. The filterability of the gypsum slurry depends on the reactor configuration and on the rock source. Highly filterable gypsum slurries are produced from Florida, Togo, and Senegal phosphates. These rocks may provide more than double the P_2O_5 from a given filter when compared to gypsum produced from other sources. However, because phosphate rock is costly to transport, the use of local, less treatable phosphates can be quite economical. Table 10.9 lists the approximate filtration design rates for these groups of commercially available phosphate rock.

Advantages of the dihydrate process versus the various, newer hemihydrate configurations are as follows:

TABLE 10.9 Filtration Design Rates for Phosphoric Acid Produced from Various Rock Sources

Rock Source	*Tons P_2O_5 Produced/Sq. Meter Active Area*
Togo Florida Senegal	7.5 to 9.0
Morocco Khouribga Western U.S.A. Kola (USSR) N. Carolina	4.5 to 7.0
Morocco Safi Algeria Tunisia	3.5 to 5.0

1. The water balance permits the use of wet rock slurry feeds. This eliminates the cost to dry the rock and a dust nuisance. This is the principal benefit that has deterred most Florida producers from using the hemi routes.
2. Dihydrate gypsum, in the case of most phosphates, filters at relatively higher P_2O_5 throughput rates than hemihydrate, comparing dihydrate filter performance at 28 percent P_2O_5 and hemihydrate at 42 percent P_2O_5.
3. Dihydrate plants are proven at huge outputs, in excess of 1500 MTPD P_2O_5, thus offering economies of scale.
4. A vast amount of operating data and experience exists on a wide variety of phosphate rocks for the dihydrate mode.
5. Capital costs for the dihydrate system can be less than for the two-stage hemi processes even when evaporation and rock grinding are considered, because of the larger reaction vessels and two stages of filtration in the hemi process.
6. The yield across the filter for the dihydrate process, generally about 96 percent of the P_2O_5 fed, is about 3 to 4 percent above the single-stage hemi processes.
7. Dihydrate process maintenance costs are substantially less than those for hemi processes due to less severe process

conditions. The on-stream factor is also higher for the average dihydrate facility.

Major Dihydrate Processes

Prayon Process. The Prayon process has evolved from the cascade system first used by the Dorr Company in the 1930s. Instead of round, steel vessels with rubber and brick lining, in the early 1960s Prayon developed a multicompartmented reinforced concrete vessel, rectangular in shape, lined with a membrane and carbon brick. The construction has proved to be exceedingly durable. A similar concrete construction has been used by Dorr/Jacobs and by Rhone Poulenc. A flowsheet for a current configuration Prayon process digestion system is shown in Fig. 10.3. The Prayon process uses vacuum cooling, a vacuum chamber through which reaction slurry is circulated to maintain a reactor temperature of 80 to 85°C. There are numerous installations of large Prayon plants in the United States although the center of process technology is Belgium.

Rhone Poulenc Process. Most of Rhone Poulenc's existing plants are of a single stirred vessel configuration. However, a two-vessel arrangement, the Diplo system, has been offered and retrofitted into a few of the previous single reactor plants.

The original Rhone Poulenc Single Tank system is an exceedingly simple plant. It is air-cooled by passing a flow of air over slurry splashers, and does not have a filter feed tank, the slurry simply overflowing the reactor to a vertical pipe on top of the filter feed pump.

The newer Diplo system is said to offer a better yield, higher P_2O_5 product acid strength, and a more filterable gypsum by providing a better concentration gradient for gypsum growth.

Most Rhone Poulenc installations are located in France and North Africa.

Nissan H Process. In this process, the rock attack is done under conditions favoring the formation of an unstable hemihydrate. The slurry then is cooled and seeded to recrystallize to gypsum at high sulfate levels, producing a gypsum low in co-precipitated P_2O_5. Yields in the 97 to 98 percent range are reported. The process, in operation since the 1960s, has been favored in many instances where the phosphogypsum can be used for wallboard

Fig. 10.3. Flow diagram of attack or reaction section of a wet process phosphoric acid plant. (*Prayon process courtesy Davy McKee Corp.*)

or other building materials. Major installations are located in Holland and Morocco.

Jacobs/Dorr-Oliver Process. Jacobs Engineering acquired the Dorrco process technology in 1974 and has carried on the annular reactor design begun by Dorr in the early 1960s. The reactor configuration is a compromise between a multicompartment system, as used in the earlier Dorr-Oliver cascade system and in the compartmented Prayon reactors, and the true single stirred vessel used by Rhone-Poulenc and Badger.

In the Jacobs system, the annulus of a large concrete tank is fitted with a series of agitators. There is baffling, but there are no walls between the agitators. High slurry recirculation rates are achieved by a combination of back-mixing and slurry pumping. Vacuum cooling normally is used.

Jacobs has installations mainly in the United States and India. One of the world's largest phosphoric acid reactors is shown in Fig. 10.4.[15] The process is offered in several configurations, high throughput, high yield, or coarse rock modes, and the reactor can be adapted to a hemihydrate mode making 42 percent P_2O_5 acid.

Badger—Isothermal Process. The Badger reactor is a draft tube mixer within a vacuum vessel. All reactants are added to this vessel, which is under vacuum, and cooling and rock digestion are achieved simultaneously. There are several relatively recent U.S. installations. While the system has low energy (power) consumption, the method is unproved on the more difficult-to-treat, high-organic, high-CO_2 phosphates.

Hemihydrate Processes for Phosphoric Acid

Methods for making higher-strength P_2O_5 acid have been known for a long time. The basic hemihydrate–dihydrate process shown in the Norsk Hydro flowsheet, Fig. 10.5,[14] is similar to the initial process attempted in 1931 at the Cominco plant at Trail, BC. The hemihydrate–dihydrate process failed there,

Fig. 10.4. Gardinier phosphoric acid plant, 1500 TPD, Jacobs process. (*Courtesy Jacobs Engineering Group, Inc.*)

Fig. 10.5. Hemi–dihydrate process. (*Courtesy Norsk-Hydro, Ltd.*)

mostly because of inadequate filters, but the Dorr dihydrate process did emerge successfully (Table 10.10).

The first large hemi plant of more modern times was the Kemira hemi–dihydrate two-stage plant at Siilinjarvi, Finland, which started operation in 1969, using the Dorr HYS process. After about four years of operation at 250 MTPD P_2O_5, its rated design, the plant was expanded but operated in the dihydrate mode. In the hemihydrate mode, filtration and recrystallization problems had plagued the plant, which operated on Kola rock.

Hemihydrate processes are available for new facilities and also for the retrofitting of existing dihydrate plants. Several conversions to higher-strength acid have been made, where the steam saved in evaporation can replace fuel. Plants that make superphosphoric acid, or where sulfuric acid plant steam is not available, are likely targets for conversion.

The single-stage hemi process is smilar to the front end of the hemi–dihydrate process. However, in this case, the hemi is sent to waste without deliberate recrystallization after washing. Yields for the single-stage process are generally below the yields of the dihydrate process, whereas yields for the two-stage hemi–dihydrate process are very high.

Table 10.11 lists the major hemihydrate plants operating in 1990.

In the family of unconventional phosphoric acid processes is the Prayon-Central Glass process, which is a dihydrate–hemihydrate process in which the gypsum initially precipitated is recrystallized to hemihydrate by adding the bulk of the H_2SO_4 required for the digestion of the rock. This hemi, then, is in the most suitable phosphogypsum form (low impurities and low water content) for use as feed to the cement–sulfuric acid (Krupp-OSW) process and for gypsum plaster. Because of the capital costs for this combined cement-sulfuric acid process and operating difficulties, there have been only a few installations.

Advantages for the hemihydrate processes compared to the conventional dihydrate processes are:

1. Energy savings due to higher product acid strength.
2. Higher P_2O_5 recovery and lower H_2SO_4 consumption for the two-stage process.

TABLE 10.10 Range of P_2O_5 Yield through Filtration as % of P_2O_5 Fed

Phosphoric Acid Process	
Hemi–dihydrate	98–99
Conventional dihydrate	95–97
Single-stage hemi	91–94

TABLE 10.11 Major Hemihydrate Installations (1990)

Owner	*Location*	*Year in Oper.*	*Process*	*Product Acid Strength*	*Rock*	*Capacity MPTD P_2O_5*
Windmill	Holland	1970	N-H(H)	42	Togo/FL	700
Windmill	Holland	1985	N-H(C)(H)	42	Togo/FL	250
Albright & Wilson	UK	1978	N-H(NDH)	42	Mor.	600
SUPRA	Sweden	1988	N-H(C)(NDH)	42	Mor./FL	360
Chinhae	Korea	1990	N-H(C)(NDH)	45	FL	250
Belledune	Canada	1986	N-H(C)(H)	39	FL	500
Arcadian	(USA)	1980	N-H(C)(H)	40	Bou Cra	600
Occidental	Florida	1974	Oxy(H)	38	FL	300
Occidental	Florida	1980	Oxy(H)	38	FL	1400
Gresik	Indonesia	1984	Nissan C(HDH)	42	Jordan	550
Nam Hae	Korea	1988	Nissan C(C)(HDH)	42	FL/Jordan	1100
Yong Nam	Korea	1989	Nissan C(C)(HDH)	42	—	400
Copebras	Brazil	1987	Nissan C(C)(HDH)	—	Brazil	450
Coop Chem.	Japan	1987	Nissan C(HDH)	—	—	230
Zhanjang	China	1993	Oxy(H)	40	China	110
Yunnan	China	1992	N-H(HDH)	45	China	210

N-H = Norsk Hydro
(C) = Conversion
(H) = Hemihydrate, single stage
(HDH) = Hemihydrate–dihydrate

3. Lower capital cost for the single stage process, and possibly lower capital cost for the two-stage process where a highly filterable hemi can be produced.
4. Hemi–dihydrate recrystallized gypsum that is relatively pure, 0.2 to 0.4 percent total P_2O_5. This makes it more suitable for a cement additive or in wallboard, etc.
5. Less rock grinding required.

Energy Savings in Phosphoric Acid. As we have just noted, the hemi processes provide energy savings due to the reduced steam consumption for evaporation. This steam, even though low-pressure, 2.0 to 3.0 kg/cm^2, has significant value for the cogeneration of power. In some cases, the hemi process also saves rock grinding power. On the other hand, dihydrate processes have the ability to use wet rock slurries and to absorb, in the process, more contaminated water resulting from rainfall collected off the waste gypsum stacks. The latter advantage may be important to maintaining a zero water balance and eliminating costly effluent treatment. The use of wet grinding and slurry feeding eliminates the fuel and electricity consumed in drying the rock.

Another energy-saving option is the use of hot water for evaporation, instead of steam. This is an alternate, energy-wise, to the higher-strength hemi processes. Hot water normally is available from the heat of absorption of the sulfuric acid plant. Strengths of 40 to 42 percent P_2O_5 can readily be achieved. A few commercial installations exist in Europe. A more expensive alternative is offered in the HRS sulfuric acid process by Monsanto, which converts the absorption heat to low-pressure steam.

Unit Operations

In addition to the reaction step discussed above, there are a number of other unit operations used in producing wet process phosphoric acid.

Calcination. Phosphate rock normally is used as a dry rock or in slurry form. However, in some cases, particularly where the raw phosphate is high in carbonaceous matter or

it is desirable to have a clean acid, the rock is calcined. Also, in a few cases, the phosphate rock is calcined, the product slaked, and free lime separated as a beneficiation step. Calcination is energy-intensive and produces a less reactive rock and in some cases a less filterable gypsum. Therefore, the use of calcination is diminishing, and in some cases has been replaced by a wet oxidation step to produce green acid. In separating calcium carbonate, flotation, where it is successful, is favored over calcination because of its lower cost.

Rock Grinding. Until 1973, most phosphate was ground dry in roller or ball mills. In that year, Agrico, at South Pierce, Florida, converted one of its dry mills to wet slurry grinding and proved that the plant water balance could manage the rock at a 65 to 68 percent solids slurry. Since that time, most U.S. installations have converted to wet grinding.

Relatively fine phosphates, such as Kola, North Florida, Senegal, and Togo, can be processed unground as dry concentrates or as dewatered beneficiated product with 12 to 15 percent moisture. In the hemi processes, somewhat coarser feed may be tolerated. For dihydrate, it is desirable to feed from $1\frac{1}{2}$ percent + 35 mesh (Tyler) to about 8 percent + 35.

Filtration. The separation of phosphogypsum or hemihydrate from its mother liquor has always been a difficult operation. The process has been subject to the formation of calcium sulfate, sodium fluosilicate, and other types of scale that clog the cloth and necessitate periodic filter washing. Filter cloth wear is severe, requiring cloth changes as often, in some cases, as two or three weeks apart.

Three types of filters have predominated over the past 20 years, the most widely used being the Bird-Prayon tilting pan filter shown in Fig. 10.6. The Ucego, a table filter with a peripheral side wall belt that leaves the filter to permit cake sluicing, has been popular worldwide since the late 1960s.

In the late 1970s and the 1980s, belt filters became more readily accepted. The belt filter

Fig. 10.6. Flow diagram of filtration section of a wet process phosphoric acid plant. (*Prayon process courtesy Davy McKee Corp.*)

has been used on phosphoric acid since the 1940s and 1950s, but in the past it was plagued by mechanical problems and materials failures. The filter, at least as offered by two vendors, has been improved substantially, but the units tend to have difficulties in containing splash and drips, and tend to dilute the product acid by water introduced as slide plate lubricant. Nevertheless, belt filters are in successful operation in many plants, particularly as supplemental filters to increase plant output.

The filtration step is a countercurrent washing using two or three washes. Usually the final wash is a contaminated pond water or a cooling loop water, thus providing for, in most cases, a zero effluent plant.

Sizes of the Bird-Prayon and Ucego filter can be very large, over 200 square meters of surface area, allowing rates up to 1600 to 2000 MTPD P_2O_5. Belt filters tend to be 100 square meters or less because of their length; however, they are relatively inexpensive because little alloy steel is used, and normally two belt filters would be less in first cost than one large tilting pan or table filter.

Evaporation. Most phosphoric acid is used at a strength of about 38 to 42 percent P_2O_5 for DAP and slurry process TSP. How-ver, many fertilizer plants produce MGA, merchant grade acid, which, at 52 to 54 percent P_2O_5 is a world-traded commodity.

Evaporation normally is done under vacuum using forced circulation via an axial flow pump. Vacuum can be accomplished by steam ejectors, vacuum pumps, or with condenser water using an entraining condenser downleg. Heat exchangers normally have been shell and tube, with graphite tubes in the United States. French practice has been to use carbon block exchangers. Some newer shell and tube exchangers have been Sanicro 28 or Hastellov G3 or G30. A flowsheet is shown in Fig. 10.7.

Typical analyses of wet process phosphoric acid made from North Carolina calcined concentrate are shown in Table 10.12.

Superphosphoric Acid

Solution fertilizers have become very popular in the United States. The principal source of

Fig. 10.7. Flow diagram of evaporation section of a wet process phosphoric acid plant. (*Courtesy Davy McKee Corp.*)

TABLE 10.12 Typical Analyses, Wet Process Phosphoric Acid Made from North Carolina Calcined Concentrate

	Weight Percentage	
Component	*Concentrated Acid*	*Superphosphoric Acid*
P_2O_5, total	53.0	69.5*
Solids	0.2	
Free water	22.0	
Fe_2O_3	1.5	2.0
Al_2O_3	0.7	1.0
F	0.6	0.3
MgO	1.1	1.3
SO_4	2.7	3.7
CaO	0.1	0.2
Sp.g, at 75°F	1.68	2.0

*About 36% of the total P_2O_5 is present as polyphosphates.
Courtesy Texasgulf, Inc.

P_2O_5 for these fertilizers is wet process superphosphoric acid containing about 70 percent P_2O_5, where 35 percent or more of the P_2O_5 is present in the polyphosphoric form. When this acid is ammoniated and diluted, the iron, aluminium, and magnesium compounds naturally present remain in solution, sequestered by the polyphosphates. Clear solutions result, and there is no clogging of the sprays used for applying the fertilizer. When ordinary phosphoric acid is neutralized with ammonia, heavy sludges form, and the resulting solution is difficult to store and apply.

Superphosphoric acid is made by the additional concentration of clarified phosphoric acid in vacuum evaporators of the falling-film or forced-circulation type. High-pressure steam or Dowtherm vapor is used for heating. Corrosion is a problem, so the equipment is made from high alloy stainless steels. The acid is shipped in special insulated tank cars to the solution fertilizer plants, which are located close to the farm areas they serve. Organic matter contributes to sludge problems in making solution fertilizers; therefore, calcination of the phosphate rock used for making the acid is advantageous. Organics also may be removed by oxidation with nitric acid or ammonium nitrate,[11] and several SPA producers have gone to such treatment to improve their product quality.

Additional details of the manufacture of superphosphoric acid, and of the triple superphosphate, are to be found in Chapter 11.

WET PROCESS ACID BY-PRODUCTS

Phosphogypsum

About 5 tons of gypsum on a dry basis are made for each ton of P_2O_5 produced in a wet process phosphoric plant. This material usually is disposed of as waste, by impounding it in old mine pits, stacking it in huge piles, or, in some cases, discharging it into very large rivers or river mouths. Phosphogypsum is sold to farmers in California for control of salt buildup in irrigated soils; a small quantity is sold to peanut farmers in the southeastern United States. However, there has been concern about the utilization of gypsum because of its low-level radioactivity in some instances. In addition, it contains fluosilicates and P_2O_5 so that utilization in building products, such as wallboard and blocks, has been limited to Europe and Japan, where natural gypsum is more costly than manufactured gypsum. In these cases, the phosphogypsum generally comes from a hemihydrate process producing a slightly purer form than natural gypsum.

Regeneration of SO_2 from gypsum has been done via the Krupp-OSW process, the most important facility formerly being the Fedmis plant in South Africa, which made about 300 TPD H_2SO_4 and a similar amount of cement. The remote location of this facility, where sulfur and cement were expensive, apparently was not enough to sustain operation, and the plant shut down in the late 1980s.

An SO_2 regeneration process, using grate calcination, under development by Davy-McKee and the Florida Institute of Phosphate Research,[12] has been pilot planted by Freeport McMoran at Uncle Sam, Louisiana. A relatively small continuous system producing

SO_2 at a rate comparable to about one ton per hour H_2SO_4 was operated for several months to produce by-product aggregate for testing as a road building material.

Much time and effort have been devoted to studying the utilization of phosphogypsum, but with only limited success. The U.S. EPA, which issued a decree in 1989 that phosphogypsum had to be put only in mined-out areas or in specific piles, has limited, at present, what can be done with the waste.

Fluorine Recovery

Phosphate rock contains about 3.5 percent fluorine, some of which is recovered as a by-product in manufacturing wet process phosphoric acid. During acidulation, the fluorine is released as hydrofluoric acid, HF, which reacts with the silica present as an impurity in the rock to form fluosilicic acid, H_2SiF_6. Some of the fluorine is lost with the gypsum as sodium or potassium fluosilicates, and some remains dissolved in the filter acid. When the acid is concentrated, much of the fluorine in the feed is boiled off, appearing as HF and silicon tetrafluoride, SiF_4, in the vapors.

Fluorine is recovered at the evaporator station by scrubbing the vapors leaving the flash chamber. The vapors pass through an entrainment separator to remove fine droplets of phosphoric acid and then into a spray tower where they are scrubbed with a weak solution of fluosilicic acid according to the reaction:

$$2HF + SiF_4 \rightarrow H_2SiF_6$$

Part of the circulating solution is continuously withdrawn as a 20 to 25 percent aqueous solution of H_2SiF_6. The solution is shipped in rubber-lined tank cars and is used for fluoridation of drinking water, the preparation of fluosilicates, and production of AlF_3. These salts find use in ceramics, pesticides, wood preservatives, concrete hardeners, and aluminium production.

Uranium Recovery

Uranium occurs in many phosphate rocks in the range of 0.005 to 0.03 percent U_3O_8. This is too low in grade for the rock to be economically processed for uranium alone. However, in the making of wet process acid, most of the uranium is put into solution and can be extracted from the filter-strength acid.

Typically, wet process phosphoric acid made from Florida rock contains about 1.0 lb of U_3O_8 per ton of P_2O_5. It is claimed that as much as 90 percent of this can be extracted. The uranium is present in solution in the acid in an oxidized state, U(VI).

Research done at the Oak Ridge National Laboratory resulted in a reduction stripping process based on using a mixture of di(2-ethylhexyl) phosphoric acid and trioctyl phosphine oxide (DEPA-TOPO) dissolved in kerosene as the extractant.

When the uranium price was about \$40/lb, uranium recovery was attractive. However, with the use of nuclear power in a holding pattern, the price of uranium in the 1980s slipped below \$20/lb. Only a few companies with long-term supply contracts were still in operation by 1990, and no new plants are likely until the uranium price rises substantially.

Animal Feed Supplements

Calcium phosphates for use in animal and poultry feeds are made from both furnace and wet process phosphoric acids. Dicalcium phosphate, $CaHPO_4$, containing 18.5 percent P, and monocalcium phosphate, $Ca(H_2PO_4)_2 \cdot H_2O$, containing 21.0 percent P, are made in large tonnages. Both grades are prepared by reacting phosphoric acid with pulverized limestone in a pug mixer. The limestone must be quite pure, and the phosphoric acid must have a low fluorine content. If 54 percent P_2O_5 wet process phosphoric acid is used, it is defluorinated first by adding diatomaceous earth and then sparging the acid with steam. An alternate method is to use wet process superphosphoric acid, which has a low fluorine content. The superphosphoric acid is hydrolyzed by diluting it with water and heating. The pug mixer product is a fine-granule, -12 mesh, which is dried and shipped in bulk to feed-mixing plants.

Purified Phosphoric Acid

High power costs and environmental problems have discouraged the installation of new phosphorus furnaces in the United States. Wet process phosphoric acid purification plants using solvent extraction are operating successfully in Europe, Japan, and the United States.

Wet process phosphoric acid made from calcined rock is preferred because organic matter in the acid is very troublesome. Acid at a concentration of 35 to 50 percent P_2O_5 is treated with various reagents and adsorbents and refiltered. Then the acid is fed to a column or a battery of mixer-settlers and extracted with a solvent such as butyl alcohol or tributyl phosphate. Generally, about three-fourths of the phosphoric acid transfers to the organic phase, leaving the impurities in the raffinate, which is sent to a fertilizer unit to recover its P_2O_5. The yield of cleaned acid can be increased by adding another mineral acid such as sulfuric acid or hydrochloric acid to the extraction step.

After washing, the phosphoric acid is stripped from the solvent with water, and the solvent is returned to the extraction section. The phosphoric acid now is quite dilute and still contains small amounts of impurities. The acid then is concentrated, and the impurities are removed by steam stripping and the addition of reagents and adsorbents followed

Fig. 10.8. Purified phosphoric acid plant for P.A. Partnership, Aurora, NC. (*Courtesy P.A. Partnership and Jacobs Engineering.*)

by filtration. The exact details of the process vary, depending upon the impurities present in the feed acid and the solvent used.

Purified wet phosphoric acid is suitable for both industrial and food-grade use. Two plants were in operation in the United States in 1991. One is pictured in Fig. 10.8. This facility, operated at Texasgulf's plant in North Carolina, is rated at 120,000 TPY P_2O_5. The second purified phosphoric acid plant is operated by Rhone Poulenc at Geismar, Louisiana. Of the North American capacity of approximately 900,000 TPY of high-purity P_2O_5 or equivalent P in 1991, about 78 percent still was furnace-derived.[13]

ENVIRONMENTAL ASPECTS

As implied earlier in this chapter, gypsum disposal is a problem that generally has defied an inexpensive solution. As a waste material, it is relatively benign, but P_2O_5, sulfate, fluorine, and other contaminants, including some heavy metals in small quantities, can leach from waste stacks into the nearby groundwater. In Florida, the underlying layers of limestone afford some protection, but the EPA will require future stacks be lined with a membrane to prevent seepage.

Although dumping of waste gypsum slurries into the ocean still is practiced in some locations outside the United States, and harmful effects are generally difficult to quantify, there is continuing pressure from environmental groups to cease dumping into rivers and the seas.

Gaseous emission from phosphoric acid plants can be handled with cool contaminated recycle cooling water to relatively low emission levels. However, to minimize cooling tower or cooling pond emission, fluorine recovery often is practiced on those streams that have a significant fluorine content.

REFERENCES

1. Leyshon, D. W., "The Origin of the Modern Dihydrate Phosphoric Acid Process—Cominco 1931," A.I.Ch.E., Orlando, FL, Mar. 1990.
2. McKelvey, V. E., "Phosphate Deposits," *Geological Survey Bulletin 1252-D*, 1967.
3. deBlinko, M. G., "Some Thoughts on the Genesis of Sedimentary Marine Phosphate Beds," *Phosphorus & Potassium*, No. 106, pp. 17–22. (Mar.–Apr. 1980).
4. Bernardi, J. P., and Hall, R. B., "Comparative Analysis of the Central Florida Phosphate District to Its Southern Extension," *Mining Engineering*, pp. 1256–1261 (Aug. 1980).
5. Northolt, A. J. G., Sheldon, R. P., Davidson, D. F., *Phosphate Deposits of the World*, Vol. 2, *Phosphate Rock Resources*, Cambridge University Press, 1990.
6. Cathcart, J. B., *World Phosphate Reserves*, American Society of Agronomy, Symposium Volume.
7. McClellan, G. H., "Mineralogy of Carbonate Fluorapatites," *The Journal of The Geologic Society of London*, **137**, 675–681 (1980).
8. Houot, R., "Beneficiation of Phosphatic Ores through Flotation: Review of Industrial Applications and Potential Developments," *International Journal of Mineral Processing*, **9**, 353–384 (1982).
9. Lawver, J. E., et al., "Beneficiation of Dolomitic Florida Phosphate Reserves," *Proc.* XIV International Mineral Processing Congress, 1982.
10. Becker, P., *Phosphates and Phosphoric Acid*, 2nd ed., Marcel Dekker, New York, 1989.
11. Leavith, Kranz, Gorman, and Stewart, "Oxidation–Reduction Process for Enhancing the Color of and Stabilizing Wet Process Phosphoric Acid," U.S. Patent 4808, 391.
12. "Third International Symposium on Phosphogypsum," sponsored by the Florida Institute of Phosphate Research, Orlando, FL, Dec. 1990.
13. "Olin and OxyChem Quit the Phosphates Business," *Chemical Week* (May 8, 1991).
14. Persson, B., et al., "Environmental Improvements and Energy Savings in the Hydro Supra Phosphoric Acid Plant," IFA Technical Conference. Vienna, Austria, Oct. 1990.
15. Leyshon, D. W., and Waters, P. S., "Energy Savings with the Jacobs-Dorrco Phosphoric Acid Process," IFA Technical Conference, Kallithea, Greece, Oct. 1982.

11

Fertilizers

George Hoffmeister*

INTRODUCTION

The growth and harvest of food and fiber crops is essential to the survival and well-being of humankind. As world population expands, or even "explodes," the challenge to increase crop production is a formidable one. The major means being employed to meet this challenge are: (1) increasing tilled acreage; (2) improvement of plant strains; (3) introduction of irrigation; (4) chemical and biological control of insects, plant diseases, and weeds; and, not the least of all, (5) introduction of or increase in fertilizer usage. It should be noted, however, that implementation of any of these methods for crop increase automatically dictates a corresponding increase in fertilizer usage. This has become especially clear during the so-called green revolution. During this period, which began in the late 1930s and extends to the present, tremendous advances in scientific crop breeding have made possible dramatic worldwide increases in the yields of corn, wheat, rice, and other staple crops, yet these increases have been sustainable only by suppling the improved varieties with increased nutrients, as fertilizer. Increasing crop acreage, introduction of irrigation, and effective use of pesticides likewise result in increased need for fertilizer in order to realize maximum yield benefits. Numerous estimates have been made in attempts to quantify the effect of fertilizers on U.S. and world food production. One of the more comprehensive indicates that approximately 37 percent of U.S. agricultural production is directly attributable to fertilizer usage.[1] However, even that statistic does not convey the full impact of fertilizer usage because without the use of fertilizer intensively farmed crop land soon would become depleted and nonproductive. For the "developing" countries of the world (as contrasted to "developed" countries) increasing the usage of fertilizer also holds great potential for reducing their dependence on imported food and for increasing human nutrition standards.

NUTRITION REQUIREMENTS OF CROPS

Twenty-one chemical elements are known to be essential to the normal development and

*Retired from National Fertilizer and Environmental Research Center, TVA, Muscle Shoals, AL.

growth of plants. Their percentage in the dry matter of the plant varies somewhat with different species of plants and with different soil and climatic conditions under which plants are grown. However, the order of magnitude, expressed as the average percentage of each element present in the dry weight of the whole plant, is roughly as shown in Table 11.1.[2]

Oxygen, carbon, and hydrogen make up the greater part of the plant structure and comprise about 95 percent of the elemental content of the plant. The plant obtains these elements from the atmosphere and from water. The remaining 18 essential elements are supplied almost exclusively by soils and fertilizers. They comprise only about 4.9 percent of the elemental content of the plant, but are absolutely essential to healthy plant growth.

Of the 18 nutrient elements, nitrogen, phosphorus, and potassium are assimilated by plants in the greater amounts, and are the ones that most often become depleted in soils. Therefore, they are referred to universally as the "primary" plant nutrients. The overriding task of the fertilizer industry is to provide fertilizers that contain these three elements in chemical forms that, when applied to the soil, are soluble in the soil solution phase and assimilable by plant roots. Most often, this means chemical forms that are water-soluble. However, in the case of phosphorus, solubility in specially specified reagent solutions (citric acid, neutral ammonium citrate, or alkaline ammonium citrate) often is accepted as indicating "availability" to plants. In the case of nitrogen, slow solubility in water may be acceptable and may in fact, in some applications be more beneficial than easy solubility.

TABLE 11.1 Elemental Content of a Healthy Plant

Element	*Amount in Whole Plant % (Dry Weight)*
Structural Elements	
Oxygen	45
Carbon	44
Hydrogen	6
Primary Nutrients	
Nitrogen	2
Phosphorus	0.5
Potassium	1.0
Secondary Nutrients	
Calcium	0.6
Magnesium	0.3
Sulfur	0.4
Micronutrients	
Boron	0.005
Chlorine	0.015
Copper	0.0001
Iron	0.020
Manganese	0.050
Molybdenum	0.0001
Zinc	0.0100
TOTAL	99.9011

*Data from *Farm Chemicals Handbook*, p. 134, Meister Publishing Company, Willoughby, OH, 1970.

In the United States, and essentially in all other countries, there is close government monitoring of the chemical compostion of all marketed fertilizers to ensure that the nutrients claimed are present in acceptable forms and in the amounts claimed. Stiff penalties are assessed for failure to meet advertised compositions. In the United States, this control is under jurisdiction of the individual states, but differences in regulations from state to state are not great.[3] Methods of sampling and chemical analysis applied by the states to fertilizers in almost all cases are developed and/or approved by the Association of Official Analytical Chemists, a nonprofit, government-chartered, scentific organization.[4]

In Table 11.1, the elements calcium, magnesium, and sulfur are shown under their usual designation as "secondary" nutrients. These elements are required by plants in moderate amounts, but in most agricultural situations are not found to be deficient in the soil. There are exceptions, however. Areas of sulfur deficiency are well recognized in Australia, some U.S. locations, and elsewhere. Correction usually is made by application of fertilizers that are naturally high in sulfur content, such as normal superphosphate and ammonium sulfate, or by special application of gypsum or elemental sulfur. A very considerable amount of sulfur enters many soils dissolved in rainwater. The origin of this sulfur

is stack gas from the burning of coal and gases from volcanic discharges. Several recent trends, however, suggest that sulfur deficiencies in the soil may increase in the future, and that sulfur fertilization thus may gain importance. One of these trends is an ongoing replacement of the sulfur-containing superphosphate and ammonium sulfate fertilizers with higher-analysis, more popular, nonsulfur fertilizers. The other trend is toward greater reduction of sulfur emissions from coal-fired power plants as a cure for "acid rain" and other environmental problems. The other two secondary nutrients, calcium and magnesium, are major chemical components of soil minerals and usually are available to plants in sufficient quantities without special supply as fertilizer. The application of lime (calcium carbonate or dolomite) to soils is a practice frequently used to regulate soil acidity, and such liming also furnishes additional calcium and magnesium.

The remaining essential plant nutrients listed in Table 11.1 are designated as "micronutrients," which reflects the fact that they are required by plants only in extremely small amounts. Nevertheless, they are essential, and deficiencies can seriously limit growth and yields.[5] Crops vary in their need for specific micronutrients. An example is boron, which is particularly required for cotton and alfalfa. There are numerous other examples. Also, there are many localized soils that are deficient in specific micronutrients. Some fertilizers, especially the older-type low-analysis superphosphates, contain significant amounts of micronutrient elements as natural impurities. However, modern scientific agriculture cannot rely on that source, especially in view of the trend toward higher-analysis fertilizers that generally contain little or no impurity usable as a micronutrient. The fertilizer industry thus includes a segment, relatively small in volume, devoted to furnishing micronutrients for fertilizer use. Often, the micronutrient compounds are incorporated in mixed fertilizers during their manufacture. As will be seen later, the now popular practice of "bulk blending" (dry mixing) fertilizer materials to "prescription" in the immediate farming locality lends itself well to the inclusion of specific amounts of specific micronutrients as required by local soils and crops.

OVERVIEW OF THE FERTILIZER INDUSTRY

For many centuries, even including ancient and medieval times, farmers have been deeply interested in improving crop yields through the addition of various mineral or organic substances to the soil. Up until the last two or three hundred years, however, the approach to the subject was highly empirical; only by accident or by trial and error was it found that applications to the soil of various organic wastes or naturally occurring mineral substances such as manure, ground bones, wood ashes, saltpeter, and gypsum dramatically improved plant growth. Still the results were not predictable; a treatment that benefited one field might have no effect—or even an adverse effect—on another.

As more and more chemical elements were identified, scientists became interested in determining the amounts and relative importance of various elements in plants. The German scientist Liebig clarified the value of elements derived from the soil in plant nutrition and stressed the necessity of replacing those elements to maintain soil fertility. He is usually credited with initiating the fertilizer industry. Liebig recognized the value of nitrogen but believed that all plants could get their nitrogen from the air, a concept that unfortunately is not true. He envisioned a fertilizer industry with nutrients such as phosphate, lime, magnesia, and potash prepared in chemical factories. In 1840, Liebig published a recommendation that pulverized animal bones be treated with sulfuric acid to make the phosphate more readily available to plants. This practice was accepted, and production of fertilizers by chemical processing thus began.[6]

Natural organic materials and various chemical by-products represented a large proportion of the total world fertilizer supply up until about the middle of the twentieth century; in the later years of the century, however, the

dependence shifted almost entirely to synthesized or chemically processed materials. Only by this means has it been possible to keep up with increasing populations, increased farm acreage, and increased plant food needs of new and improved crop varieties. Today, the fertilizer industry utilizes many facets of highly sophisticated chemistry and engineering in the manufacture of fertilizers. The procuring and handling of raw materials and the distribution and marketing of products also involve the latest technology and many innovations. In addition, the agronomic aspects of fertilizer usage engage the efforts of many individuals and organizations. Both private companies and state and national governments operate laboratories to serve farmers with analysis of soils, prescription of fertilizers, and other services. In the United States, extensive research and development on improving fertilizer production and usage has been carried out by the U.S. Department of Agriculture and by the National Fertilizer Development Center of the U.S. Tennessee Valley Authority, as well as by numerous private companies.

The fertilizer industry is a large-volume one, usually classified among the "smokestack" industries. The magnitude of this industry is apparent from Fig. 11.1, which shows the yearly U.S. consumption of primary plant nutrients (N, P_2O_5, and K_2O) since 1965. For the most recent year shown (July 1, 1989, through June 30, 1990), U.S. total consumption of N, P_2O_5, and K_2O was 20.6 million short tons. The average nutrient content of all fertilizers for that period was 45.9 percent; so it is apparent that some 45 million tons of fertilizer was produced and handled by the industry during that year. In general, the industry includes a number of "basic producers," each of which typically concentrates on producing large amounts of single-nutrient or bi-nutrient fertilizer products, usually at locations near raw materials. Numerous other components of the industry concentrate on locating close to local farmers and distributing the basic products either in the form received from basic producers or after various secondary processing operations such as granulation, blending, or conversion to fluids.

Fig. 11.1. U.S. consumption of plant nutrients in fertilizers. (*Courtesy TVA.*)

NITROGEN FERTILIZERS

Nitrogen, one of the three primary nutrients required by growing plants, comprises 79 percent (by volume) of the earth's atmosphere. Above each acre of the earth, the atmosphere contains about 36,000 tons of nitrogen,[7] essentially all in elemental form. Ironically, however, only a very limited number of plant varieties, chiefly legumes, can utilize this nitrogen directly from the air. For most plants, including such important cereal crops as rice, wheat, and corn, the nitrogen uptake must be from "fixed nitrogen" (nitrogen in chemical compound form) dissolved in the soil solution. Lightning discharges, and perhaps other atmospheric phenomena, convert small amounts of atmospheric nitrogen to nitrogen oxides, which then enter the soil dissolved in rainwater. However, the resultant natural nitrogen content of soils is normally very low (less than 0.3%), and it becomes rapidly depleted if cropping is carried out without replenishment.[8]

Prior to about 1945, the use of chemical nitrogen fertilizers was considered of only minor importance, and was practiced on a relatively limited scale. The more widely accepted method of supplying nitrogen was through the application of manures (organic wastes) and the use of crop rotation. Crop rotation usually involved growth of a nitrogen-fixing legume crop, such as peas, clover, or alfalfa, which then, with its content of "fixed" nitrogen, was plowed into the soil to furnish nitrogen for a subsequent nonlegume crop, for example, corn. This method is an effective one and gives the added bonus of improving the physical condition of the soil through the introduction of organic matter. However, such crop rotation is relatively inefficient in regard to land usage and labor requirement. Modern, high-yield agriculture cannot, under most conditions, be sustained by such a system.

Miscellaneous Low-Volume Nitrogen Fertilizers

Natural Organics. Animal and human wastes have long been used as fertilizers, especially in Europe and Asia. Even some American Indians are said to have planted a dead fish in each corn hill to increase yields. Such materials contain small percentages of nitrogen and other plant nutrients that are assimilable by plants. Today, the use of raw sewage on crops persists in Asia and Europe, but volumewise is not of great significance. In the United States, the use of raw sewage is not considered acceptable, but some municipal sewage plants employ special sterilization and drying processes that yield acceptable fertilizers of low nitrogen content. The volume of such output is small, however. Other organic materials traditionally employed as fertilizers of usable nitrogen content include guano (deposits of accumulated bird droppings), fish meal, and packinghouse wastes including bone meal and dried blood. However, the cumulative importance of all such natural nitrogen sources in modern agriculture is minor. Less than one percent of the total fertilizer nitrogen now used comes from such sources. These products, for the most part, are not chemically altered prior to use. Processing is mainly physical in nature, to improve handling and distribution properties.

Mineral Nitrogen. The only nitrogen-containing mineral that is mined commercially for fertilizer use is saltpeter (sodium nitrate, $NaNO_3$). An impure form of sodium nitrate known as "coliche" is found in large, rocklike deposits in a relatively dry desert area of Chile, and these deposits have long been mined to provide natural sodium nitrate for the fertilizer market. The ore is blasted from open mines, sized, and leached with water in vats. Evaporation of the resultant solution yields the commercial crystalline fertilizer product known widely as "Chilean nitrate," "nitrate of soda," or among some farmers simply "sody." Its nitrogen content is 16 percent. This product first was imported and marketed in the United States in 1830; so it is one of the oldest fertilizers on the American market. Sodium nitrate now, however, is a relatively insignificant component of the world's nitrogen fertilizer market, having been supplanted by synthetic nitrogen compounds. In the year that ended

June 30, 1990, sodium nitrate use as fertilizer in the United States was equivalent to only 8400 short tons of nitrogen, out of a total usage of over 11 million tons of nitrogen in all fertilizers.[9]

By-product Ammonia. Coke, an essential ingredient in the smelting of iron ore and in some other metallurgical processes, is produced by the heating of coal in an essentially oxygen-free atmosphere. This heating volatilizes numerous hydrocarbon components of the coal and results in decomposition of some of these compounds. The nitrogen content of coals normally is about one percent, and approximately half of this is volatilized as ammonia gas in the coking process. Recovery of this ammonia has long been effected by passing the coke-oven effluent gas through scrubber-crystallizers containing sulfuric acid. The resultant crystalline by-product ammonium sulfate is an excellent fertilizer source of both nitrogen and the secondary plant nutrient sulfur. The nominal nitrogen content of the product is 20 percent, and the sulfur content is 24 percent. Prior to the introduction of synthetic ammonia production processes (1913–25), coke-oven ammonia was a major source of fertilizer nitrogen; however, its production was very small when compared to today's demand for fertilizer nitrogen. Today, ammonia production from coke ovens, although still carried out, is insignificant in volume compared with that from ammonia synthesis processes.

Nitrogen Fertilizers from Synthetic Ammonia

Over 95 percent of the 11 million tons of nitrogen supplied to U.S. farmers yearly in commercial fertilizers originates as synthetic ammonia made from air, water, and either a petroleum-based hydrocarbon or coal. The development of a practical ammonia synthesis process in the early years of the twentieth century was a profound scientific achievement of great social significance, in view of the subsequent dependence of the world on fertilizer for support of its growing population. Both the historical background and detailed technical descriptions of modern ammonia synthesis processes are covered in Chapter 28. These processes all are refinements of the original one conceived and developed by Fritz Haber and Carl Bosch in Germany during the period 1904–13. Basically, a gaseous mixture of nitrogen and hydrogen, in proper proportions to form ammonia, is compressed to very high pressure in the presence of an activated iron catalyst. Ammonia forms and is removed by cooling and condensation. The nitrogen feed is obtained from air by any of several routes that remove the oxygen. Likewise, there are several routes by which hydrogen feed is obtained, most of which involve decomposition of water by reaction with a carbon source such as natural gas, naphtha, or coke. The preparation and the purification of the synthesis gas and the synthesis itself are highly sophisticated modern processes that require great skill and know-how in design, construction, and operation of plants.

There are many processing routes by which synthetic ammonia subsequently finds its way into finished fertilizers. The major routes are outlined graphically in Fig. 11.2 and are discussed in some detail below. Figure 11.2 also shows the approximate percentage of total U.S. fertilizer nitrogen consumption currently furnished by each type of finished fertilizer.

Anhydrous Ammonia. Anhydrous ammonia, without further processing, is an excellent nitrogen fertilizer when properly injected in gaseous form below the surface of many soils. Such direct injection first came into farm use in the early 1940s and rapidly became accepted. The procedure today accounts for approximately one-third of total U.S. nitrogen fertilization (Fig. 11.2). This proportion of the nitrogen fertilizer market is greater than that claimed by any other finished fertilizer product. Anhydrous ammonia is also the major nitrogen fertilizer in Canada, Denmark, and Mexico, but in some other countries it is, for various reasons, a relatively minor nitrogen source.[10]

Anhydrous ammonia injection is popular for economic reasons. Ammonia is naturally

Fig. 11.2. Major routes of synthetic ammonia into finished fertilizers. (*Courtesy TVA.*)

the cheapest form of fixed nitrogen at the point of production, as no further processing is required. Also, the nitrogen content of anhydrous ammonia (88.2% N) is almost twice that of the next most concentrated nitrogen fertilizer (urea: 46% N); this high nitrogen content reduces shipping and handling weight proportionally.

From production to field application, ammonia is handled in liquefied form. Because ammonia is a gas at normal atmospheric pressure and temperatures, its maintenance in liquid form requires either pressurization or refrigeration. Although such handling requires specialized equipment and considerable safety precautions,[11] there are advantages and economies inherent in the handling of a fluid rather than a solid.

The vapour pressure of anhydrous ammonia as a function of temperature is shown in Fig.

Fig. 11.3. Vapor pressure of anhydrous ammonia vs. temperature.

11.3. From this figure it can be seen that to maintain ammonia in liquid form at atmospheric pressure it is necessary to refrigerate it to at least −28°F (−33°C). In storage and handling without refrigeration, rather high pressures must be contained when atmospheric temperatures reach high levels. For example, a pressure of 265 psi, which, for safety reasons, has been set as the upper limit allowable for storage vessels at retail locations,[11] will maintain the liquid form only up to a temperature of 115°F. Safety pop-off valves are set to vent ammonia at higher pressures.

Facilities for handling liquid ammonia from points of production to points of use are highly developed in the United States. A considerable amount of ammonia is moved hundreds of miles through interstate pipelines. Major pipelines run from points of production in Texas, Oklahoma, and Louisiana oil fields all the way to the intensively farmed midwestern area of the United States. Storage at the points of production and at large terminals sometimes is in pressurized vessels (spheres), but more often it is at atmospheric pressure in large, insulated, refrigerated vessels. Refrigeration is effected simply by allowing vaporization of some of the contained ammonia, compressing the resulting vapor, and recondensing it in a water-cooled condenser. Barges with similar self-sustaining refrigeration systems commonly move large amounts of anhydrous ammonia on rivers and on the high seas. Also, specially designed railroad tank cars ranging in capacity from 24 to 73 tons are used widely for ammonia transport. These cars typically are non-refrigerated, instead depending on pressure retention to maintain the liquid form. Large pressurized tank-trucks also are used. At retail depots in farming areas, storage normally is in pressure vessels, often with some compressor capability to recondense vapor and to facilitate transfers. Movement of ammonia to farm fields is via wheeled pressure tanks known as "nurse" tanks. Application equipment, usually tractor-drawn, is equipped with pressure tanks that are filled by transfer from nurse tanks.

Because anhydrous ammonia is gaseous at atmospheric conditions, for fertilizer use it must be injected below the surface of the soil to ensure absorption by the soil solution. Most applicators are equipped with a metering system and a series of soil-penetrating "knives"; the metered ammonia is piped through tubing down the trailing edge of each knife to a single opening at the deepest extremity. A penetration of 15 to 25 centimeters is typical.[10] Retention of ammonia in the soil is best when abundant soil moisture is present.

Aqua Ammonia. Ammonia in simple water solution is popular as a nitrogen fertilizer in some locations, but overall it claims less than one percent of the total nitrogen fertilizer market (Fig. 11.2). This product is made from anhydrous ammonia in units referred to as "converters." These units simply provide cooling as ammonia and water are combined

in metered proportions. Aqua ammonia normally contains 20 percent nitrogen (23% NH_3). Its chief advantage over anhydrous ammonia is its low vapour pressure; this allows handling in nonpressure equipment, which simplifies design and reduces safety hazards. The chief disadvantage of aqua ammonia as a fertilizer is its relatively low nitrogen content, which limits its shipping range and increases field handling requirements. As with anhydrous ammonia, the preferred application method is knifing below the soil surface to limit volatilization loss; however, a relatively shallow injection depth of only 3 to 5 inches can be used.

Ammonium Nitrate and Urea. Almost half of the total fertilizer nitrogen application in the United States is as ammonium nitrate or urea, in either solid or solution form. The breakdown for the year that ended June 30, 1990 (Fig. 11.2) is: ammonium nitrate solid 7 percent; urea solid, 21 percent; and urea–ammonium nitrate solutions, 20 percent. The history and the technology of ammonium nitrate and urea production from synthetic ammonia are covered in detail in Chapter 28. The production of nitric acid from synthetic ammonia, which is a prerequisite to the production of ammonium nitrate, also is covered completely in Chapter 28. The production plants for all these materials are, like synthetic ammonia plants, highly sophisticated, involving all facets of modern engineering.

The production processes for both ammonium nitrate and urea yield their products first in the form of saturated solutions. Typical ammonium nitrate solution concentration is 75 to 85 percent (26–30% N) at 105 to 170°F, and typical urea solution concentration is about 75 percent (34% N) at 200°F. To produce the solid forms of these products, the solutions first must be concentrated extremely low water contents (1–2% at 350–285°F) and then fed to prilling or granulation equipment. The resultant solid products (Fig. 11.4) are highly acceptable to farmers as high-nitrogen fertilizers and to "bulk blenders" as nitrogen sources in the formulation of dry-blend mixed fertilizers, as will be described later. The nitrogen content of solid ammonium nitrate is 34 percent, and that of urea is 46 percent.

The concentration and the solidification of ammonium nitrate and urea solutions to form solid fertilizers require considerable expenditure of energy and processing expense; so it is attractive to use the solutions in the formulation of liquid fertilizers, without processing them into solid form. However, neither ammonium nitrate nor urea alone is sufficiently soluble at ambient temperatures to be attractive for direct use as liquid fertilizer. With either compound alone, a solution with a salting-out temperature of 32°F would contain only about 18 percent nitrogen. Fortunately, however, the joint solubility of the two compounds in water is much more favorable than this; so the liquid nitrogen solutions marketed for fertilizer use almost invariably are joint solutions of ammonium nitrate and urea. Typical commercial solutions contain 28 to 32 percent nitrogen, with compositions and characteristics as shown in Table 11.2.

The high-volume use of these solutions is a result of their economy of production and the safety and convenience of their handling and application. The use of nonpressure equipment and surface application (in contrast to knifing) are normal. Nitrogen solutions of other compositions also are prepared commercially, but are used primarily as intermediates in the preparation of solid fertilizers. Such solutions usually consist of various combinations of water, ammonia, ammonium nitrate, and/or urea. Because of the content of free ammonia, most of these solutions require pressurized handling.

Ammonium Sulfate. Although an excellent nitrogen–sulfur fertilizer with good physical properties, ammonium sulfate is now outranked economically by ammonium nitrate and urea. Only about 4 percent of U.S. fertilizer nitrogen is furnished as ammonium sulfate. The chief reasons for the unpopularity of ammonium sulfate are its relatively low nitrogen content (21%) and its relatively low solubility in liquid fertilizers. At most U.S. locations, it is uneconomical to produce ammonium sulfate

Fig. 11.4. Some major fertilizers in the prilled or granular form preferred by some farmers. (*Courtesy H. O. Hester.*)

TABLE 11.2 Physical and Chemical Characteristics of Urea–Ammonium Nitrate Nonpressure Nitrogen Solutions*

Grade, % N	*28*	*30*	*32*
Composition by weight			
Ammonium nitrate, %	40.1	42.2	43.3
Urea, %	30.0	32.7	35.4
Water, %	29.9	25.1	20.3
Specific gravity, 15.6°C (60°F)	1.283	1.303	1.32
Salt-out temperature, °C (°F)	−18 (1)	−10 (+14)	−2 (+28)

*Hignett, T. P. (Ed.), *Fertilizer Manual*, International Fertilizer Development Center, Muscle Shoals, AL (1978)

from synthetic ammonia and virgin sulfuric acid. As a result, most of the ammonium sulfate available for fertilizer use is a by-product of other processes. Its production by scrubbing ammonia-containing coke-oven effluent gas with sulfuric acid already has been mentioned. A larger source now is by-product ammonium sulfate solution from the production of caprolactum and acrylonitrile, which are synthetic fiber intermediates. This by-product solution normally is fed to steam-heated evaporator-crystallizers to yield the solid ammonium sulfate. Additional production of by-product ammonium sulfate is obtained from "spent" sulfuric acid, by reaction with synthetic ammonia. In this case, the production of ammonium sulfate is a method of disposing of acid that already has been partially spent in other processes.

By-product ammonium sulfate produced by any of the above-mentioned methods is usually in the form of crystals that are 90 percent or more finer than 16 mesh, and thus are much smaller than the size preferred for present-day finished fertilizers. This small size is not undesirable if the material is to be used as feed to a fertilizer graulation plant, but it is undesirable for bulk-blending (dry mixing) or direct field application. It is possible to produce suitably large crystals, but only at the expense of greatly reducing crystalizer capacity. Both compaction and graulation processes have been employed with some success on a limited scale to convert the small-crystal material to larger-size product.[12,13]

Ammonia in Multi-nutrient Fertilizers. The ammonia-based fertilizers discussed thus far all contain only one of the three major plant nutrients, namely, nitrogen. Referring again to Fig. 11.2, one may see that ammonia is the source of fertilizer nitrogen also, either directly or via nitrogen solids or solutions, in multi-nutrient fertilizers. These are fertilizers that contain two or three of the major plant nutrients–nitrogen, phosphorus, and potassium. Ammonium phosphates, both mono- and di-, are phosphate fertilizer products that also carry nitrogen. (These will be discussed later, as major suppliers of phosphate.) As suppliers of nitrogen, the ammonium phosphates furnish only about 9 percent of all U.S. fertilizer nitrogen. Other multi-nutrient solid fertilizers, together with multi-nutrient fluid fertilizers, furnish an additional 5 percent of the total U.S. fertilizer nitrogen. (Production of these fertilizers also will be described later.) The source of their nitrogen is, in all cases, ammonia and/or ammonia-based solids or solutions. Nitric phosphates are suppliers of nitrogen through their content of nitric acid and ammonia. They are not presently produced, or used to any significant extent, in the United States, but are produced in Europe and some other locations. (Their production too will be discussed briefly later.)

PHOSPHATE FERTILIZERS

Phosphorus intake is essential to the metabolism of both plants and animals. In the case of plants, the entire uptake is from the soil solution via root absorption. Absorption occurs only when the phosphorus is present in the soil solution in the form of $H_2PO_4^-$, HPO_4^{--}, or PO_4^{---}, ions.[8] Native soils have an extremely low phosphorus content; in the United States, for example, the range is from 0.0 to 0.3 percent P_2O_5 equivalent.[8] Furthermore, a large proportion of the soil phosphorus is present in chemical forms that are not "available" (assimilable), or are only very slowly available, to crops. Cropping of native soils without phosphate fertilization soon depletes the supply of phosphorus and renders the soil barren. In times before fertilizers became readily available, it was not uncommon to "wear out" a farm by repeated cropping. Crop rotation, which was mentioned earlier as a useful method of converting atmospheric nitrogen to available soil nitrogen, is ineffective as a means of phosphorus fertilization. In fact, the nitrogen-converting legumes generally are voracious consumers of phosphate.

Modern, high-yield agriculture is dependent on regular fertilization with phosphorus compounds that either are immediately soluble

in the soil solution or that become soluble at a rate sufficient to supply the crop. A factor to contend with in phosphate fertilization is soil "fixation" of phosphorus; that is, the characteristic of many soils to convert at least a portion of applied phosphate to chemical forms that are unavailable to plants. The fixation tendency varies with soil type, and generally is highest in soils of high clay content. The overall result of fixation is that considerably more phosphate must be applied as fertilizer than is removed with crops. Some estimates indicate that, nationwide, as much as 60 percent of applied phosphate may be lost to fixation.

Natural Organic Phosphate Fertilizers

In animal metabolism, phosphorus concentrates in the bones and, in the case of crustacea, in the shells. Manures and human wastes, previously mentioned as effective but low-volume sources of fertilizer nitrogen, are much less efficient sources of phosphorus than of nitrogen. The phosphorus content of such products is normally only 1.3 to 3 percent, which is only 25 to 50 percent as great as their nitrogen content. Bird droppings, including chicken manure and guano deposits, are somewhat higher in phosphorus content (7% P), but are available in only relatively insignificant amounts overall.

Raw animal bones normally contain 8 to 10 percent phosphorus (20–25% P_2O_5); thus they are a relatively rich source. The commerical grinding of bones for fertilizer use began in Europe in the early nineteenth century.[6] The practice persists today in the production of bone meal, but only in very low volume, chiefly to furnish slowly available nutrients for greenhouse and house plants. In about the year 1830, it was found that pretreatment of bone meal with dilute sulfuric acid greatly enhanced the "availability" of the phosphorus to plants, and the practice became common. The initiation of this practice now is regarded to have had considerable historic significance as the apparent beginning of the chemical fertilizer industry. The supply of bones being very limited, it soon was found (about 1840) that a similar acid treatment of mined phosphate ore rendered the phosphorus "available" as a fertilizer. Thus was born the still important "superphosphate" industry, and the supplanting of the very limited organic phosphorus sources by widely available mineral sources was begun. Today, organic sources of fertilizer phosphorus account for less than one percent of total worldwide consumption.

Fertilizers from Mineral Phosphates

Essentially all fertilizer phosphorus now is derived from mined ores. (The occurrence, mineral characteristics, mining, and beneficiation of major phosphate ores were described in some detail in Chapter 10.) Worldwide, about 85 percent of the mined phosphate eventually finds its way into fertilizer.[10] Estimates of the world reserves of ore vary widely, but even the most conservative estimates indicate a sufficiency for hundreds of years at expected consumption levels. Supply problems of the immediate future will relate chiefly to exhaustion of the better ores, with the result that ores of lower grades and higher impurity contents will have to be processed.

Major routes of mineral phosphate (phosphate rock) into finished fertilizers are outlined in Fig. 11.5 and are discussed below. It is obvious from the figure that although phosphate rock is used directly in several major fertilizer production processes (ordinary superphosphate, nitric phosphates), most important processes require that the rock phosphate first be converted to phosphoric acid (H_3PO_4). Phosphoric acid production, then, is a very significant component of the phosphate fertilizer industry. The processes used for the production of fertilizer-grade phosphoric acid are known collectively as "wet" processes, and involve, initially, dissolution of the rock in sulfuric acid. (The highly specialized technology of wet-process acid production was discussed in Chapter 10.)

Phosphate Rock for Direct Application. Application of finely pulverized raw phosphate

Fig. 11.5. Major routes of mineral phosphate into finished fertilizers. (*Courtesy TVA.*)

rock directly to the soil has been utilized almost from the beginning of fertilization practice. Although such direct use eliminates the need for most processing, there are major drawbacks. The ores are of low phosphorus content in comparison to most manufactured fertilizers, which increases shipping cost. Also, the very fine grinds that are required to encourage solubility are difficult to handle and to apply. The major drawback, however, is agronomic. The agronomic effectiveness of raw rock depends upon the chemical and mineralogical nature of the particular rock, the pH and other characteristics of the soil, and the crop produced. Researchers at TVA have characterized phosphate rocks from a large number of sources as to their reactivity.[14] Phosphate rock from North Carolina (U.S.A.) and Gafsa (Tunisia) are at the top, and are about equal in reactivity and suitability for

direct application. Other phosphate rocks that are reasonably reactive and are marketed for this purpose include some types from Morocco, Israel, and a few other locations. In 1974, a world total of about 3.8 million metric tons of phosphate rock was used for direct application. In the United States, however, less than 10,000 tons of P_2O_5 per year currently is applied as direct application rock. This represents less than one percent of total P_2O_5 fertilization.

Ordinary Superphosphate (OSP). In 1842, Lawes in England followed up on the pioneering work of Liebig and received a patent on the use of sulfuric acid in reaction with raw phosphate material to produce "superphosphate" fertilizer. An industry grew slowly, and in 1862 about 150,000 tons of what later would be referred to as "ordinary" (OSP) or "normal" (NSP) superphosphate (16–18% P_2O_5) was produced in England. By 1870, there were 70 ordinary superphosphate plants in the United Kingdom and 7 in the Charleston, South Carolina, area of the United States.[15] In 1888, commercial shipment of phosphate rock from Florida was initiated, and a major mining industry followed in that state, which continues today. As late as 1955, OSP supplied over 60 percent of the world's phosphate fertilization and was the major phosphate fertilizer in the United States. Rock from Florida or South Carolina was shipped throughout the country to local OSP production facilities. These facilities usually included the capability of blending the OSP with nitrogen and potash materials and bagging the resultant mixtures.

In recent years, as indicated in Figure 11.6, U.S. production of OSP (shown as NSP), to a large extent, has given way to production of the higher-analysis "triple superphosphate" and ammonium phosphates. The U.S. production of OSP in the year that ended June 30, 1990 was equivalent to only 66,000 tons of P_2O_5 for direct application, plus perhaps an additional 40,000 tons for use in the manufacture of mixed fertilizers. This total production represents only about one percent of the U.S. fertilizer P_2O_5 usage. OSP retains greater importance than this in some countries. For example, OSP remains popular in Australia, at least in part because its sulfur content is particularly needed there.

OSP is the simplest, as well as the oldest, of manufactured phosphate fertilizers. Its manufacture consists of reacting pulverized phosphate rock with sulfuric acid in comparatively simple equipment, followed by sufficient aging (curing) to complete the reaction and improve the physical condition. The acid–rock reaction converts the water-insoluble apatite structure of the rock to soluble monocalcium phosphate. When produced from high-grade rock of 30 to 32 percent P_2O_5 content, OSP contains about 20 percent available P_2O_5; product made in earlier years from lower-grade rocks contained only 14 to 18 percent P_2O_5. OPS can be used for direct application, for bulk blending, or in the production of granular NPK fertilizers. If the product is to be used for direct application or bulk blending, it usually is granulated to improve handling properties.

The main equipment for manufacturing OSP consists of a mixer to bring together sulfuric acid and phosphate rock, followed by a den, and, if granulation is used, some suitable type of equipment for granulation with steam or water. Early plants used batch-type mixers for acid–rock mixing, but these mixers mostly have been replaced by continuous types. Several types of continuous mixers have been employed; of these, the lowest-cost, simplest, and apparently most satisfactory one is the TVA cone mixer, which has no moving parts (see Fig. 11.7). Mixing is accomplished by the swirling action of the acid. Short, single-shaft or double-shaft pug mills also are used for continuous mixing. The mixers discharge into a den, the purpose of which is to allow time for the acid–rock reaction to progress and the physical condition to reach a dry, friable state. The dens can be of either the batch or the continuous type. A slat-conveyor continuous den, commonly referred to as a Broadfield den, is used in many continuous systems.[16] A diagram of a typical OSP plant employing a cone mixer and a

Calendar Year	--Superphosphate-- Normal	Triple	------Multiple-Nutrient Materials------ DAP	MAP	Other*	Total	Total
			(thousand short tons of P205)				
1965	1,113	1,466	...	...	...	1,252	3,830
1966	1,138	1,696	...	...	...	1,615	4,450
1967	1,184	1,481	...	...	...	2,030	4,695
1968	938	1,419	...	...	...	1,839	4,196
1969	807	1,354	...	...	...	2,132	4,292
1970	670	1,474	...	...	...	2,452	4,596
1971	626	1,503	...	...	...	2,863	4,992
1972	677	1,659	...	...	...	3,147	5,482
1973	619	1,693	...	...	...	3,226	5,538
1974	698	1,719	2,100	...	626	2,950	5,367
1975	484	1,678	2,655	...	582	3,411	5,573
1976	383	1,595	2,876	...	677	3,847	5,824
1977	340	1,791	3,455	...	779	4,568	6,699
1978	291	1,820	3,936	...	807	5,065	7,176
1979	353	1,842	4,257	...	819	5,468	7,662
1980	425	1,693	4,972	...	854	6,191	8,309
1981	237	1,491	4,076	529	576	5,181	6,909
1982	139	1,065	3,681	484	237	4,402	5,606
1983	122	1,246	4,782	710	197	5,689	7,056
1984	127	1,124	5,804	844	134	6,781	8,032
1985	100	1,190	5,340	911	112	6,363	7,653
1986	65	972	4,222	741	109	5,071	6,108
1987	64	956	5,017	963	105	6,085	7,105
1988	86	976	5,450	1,084	82	6,615	7,677
1989	67	832	6,175	1,061	102	7,338	8,237
1990	66	929	6,427	1,119	90	7,636	8,631

*Includes MAP 1974-1980.

Fig. 11.6. Production of phosphate fertilizer materials in the United States. (*Source: ESDC, "Inorganic Fertilizer Materials and Related Products," Series M28B, monthly and annual reports.*)

Fig. 11.7. Typical cone mixer for production of 25 to 30 tons/hr of normal superphosphate. (*Courtesy TVA.*)

continuous den (slat-conveyor) is shown in Fig. 11.8.

The reaction of phosphate rock with sulfuric acid to produce OSP can be approximated in chemical stoichiometric terms. However, in general practice the proportioning usually is based on a simpler relationship of about 0.6 lb of sulfuric acid (100% H_2SO_4 basis) per pound of phosphate rock (30–32% P_2O_5). The phosphate rock usually is pulverized to about 90 percent −100 mesh and 70 percent −200 mesh.

Gases that are released while the superphosphate is solidifying (setting) cause the mass in the den to become porous and friable so that it can be "cut" (disintegrated) and handled readily. OSP made from typical rock will "set" in 40 to 50 minutes in a continuous den, whereas the set time in a batch den is 1.5 to 2 hours. The superphosphate usually is held in storage piles (cured) for 4 to 6 weeks in order to obtain better handling properties and to allow the chemical reactions to continue. The usual grade of OSP made from Florida rock is 20 percent available P_2O_5. A

Fig. 11.8. Continuous process for the manufacture of normal superphosphate. (*Courtesy TVA.*)

typical analysis is shown below.

Analysis, % by Weight								
P_2O_5								
Total	*Available*	*H_2O Soluble*	*Free CaO*	*Acid*	*SO_4*	*F*	*R_2O_3*	*MgO*
20.2	19.8	18.0	28.1	3.7	29.7	1.6	1.6	0.15

Because of the low phosphate analysis of OSP (20% P_2O_5), economics favor shipping the rock (32% P_2O_5) to local plants and producing the superphosphate there. In present U.S. practice, a majority of the producers use the locally produced OSP in formulations for granular NPK fertilizers made in a TVA ammoniation-granulation process. The maximum amount of OSP possible is used in the NPK formulations, as it usually is the lowest-cost form of P_2O_5; however, because of the relatively low P_2O_5 content of OSP, not much can be used in high-analysis NPK formulations. OSP will readily absorb about 6 lb of ammonia per 20 lb of P_2O_5 during the ammoniation-granulation process. Production of granular OSP can be carried out by starting with either cured or ex-den OSP. The OSP is fed to a rotary drum or pan granulator along with a water spray and/or steam. The granulator is followed by a dryer and screening (sizing) system. Fines and crushed oversize material are fed back to the granulator.

Triple Superphosphate (TSP). Triple superphosphate is made by acidulation of phosphate rock with phosphoric acid, using equipment and processes similar to those for OSP. As in OSP production, the acid–rock reaction converts the water-insoluble apatite structure of the rock to water-soluble monocalcium phosphate. TSP, with its relatively high P_2O_5 content of 45 to 46 percent, did not appear on the scene in any appreciable quantity until wet-process phosphoric acid (see Chapter 10) was produced commercially. TVA carried out demonstration-scale production of triple superphosphate in the late 1930s by using electric-furnace phosphoric acid. Extensive agronomic testing and market development, through the use of large tonnages of this TSP by farmers and fertilizer manufacturers in TVA demonstration programs, led to the rapid acceptance, of TSP. Producers of phosphate rock in the United States and other countries moved into production of wet-process acid and TSP. Logistics favored production of the higher-analysis TSP (46% P_2O_5 vs. 32% P_2O_5 in phosphate rock) near the source of the phosphate rock, and shipment of the TSP to distribution facilities or blending plants near the markets. Today, the largest producers of TSP are located adjacent to phosphate mining operations in Florida, North Carolina, and the western United States. Since its introduction, TSP has steadily displaced OSP in the marketplace, as shown in Fig. 11.6. In the year that ended June 30, 1990, TSP provided 11 percent of U.S. phosphate fertilization, as compared to about one percent for OSP (Fig. 11.5).

The TVA-developed cone mixer has been used almost universally in the production of nongranular TSP. Because the "set time" for TSP is only 14 to 20 minutes, as compared to 40 to 50 minutes for OSP, a simple, cupped conveyor belt usually is used to hold the acidulate until the TSP solidifies; this is instead of the slat-type den used for OSP. With a belt about 5 feet wide and 100 feet long, the production rate usually is 40 to 50 tons/hr. A flow diagram for TSP production by the cone mixer–"wet-belt" system is shown in Fig. 11.9.

Proportioning for TSP typically is 2.4 to 2.5 lb of P_2O_5 from acid for each pound of P_2O_5 from rock. Nongranular TSP usually is cured 4 to 6 weeks prior to shipment or use at the site. Typical chemical analysis of TSP

Fig. 11.9. Continuous system for manufacture of nongranular triple superphosphate. (*Courtesy TVA.*)

made with Florida rock is shown below:

Analysis, % by Weight								
P_2O_5								
Total	*Available*	*H_2O Soluble*	*Free Acid*	*CaO*	*R_2O_3*	*MgO*	*F*	*H_2O*
46.9	46.3	42.0	3.4	19.3	3.1	0.5	2.7	4.5

The use of TSP in granular NPK fertilizer formulations, together with or in place of OSP, has allowed the production of higher-analysis grades of granular NPK fertilizers, such as 13-13-13 instead of 10-10-10. The TSP can readily be ammoniated to about 3.5 lb of ammonia per 20 lb of P_2O_5 during the granulation process.

TSP is produced in large quantities in granular form for use in direct application and for use in bulk blends. In some processes, cured or ex-den TSP, like OSP, is granulated in a rotary drum or pan granulator using steam and water to promote granulation.[10] A slurry-type granulation process, as outlined in Fig. 11.10, is used in the United States and some other countries. Pulverized phosphate rock is treated with wet-process phosphoric acid in a two-stage reaction system, and the slurry is sprayed into a pug mill or rotary drum, where it is layered on recycled undersize and crushed oversize at a ratio of 1 lb per 10 to 12 lb of recycle. Product granules (45% P_2O_5) are quite spherical and dense. The lower grade of the granular product, as compared to that of cured nongranular TSP, results from the need to decrease the acidulation ratio from the usual 2.45 lb of acid P_2O_5 per pound of rock P_2O_5 to about 2.25 in order to control the free-acid content and to minimize stickiness. For production rates higher than about 25 tons/hr, a rotary drum

Fig. 11.10. Slurry process for production of granular triple superphosphate. (*Courtesy TVA.*)

usually is used instead of a pug mill. Most of the granular TSP produced in the United States is used in bulk blending, as is discussed later.

Nitric Phosphate. Fertilizers that are referred to as nitric phosphate or nitrophosphate are produced by acidulation of phosphate rock with nitric acid or with mixtures of nitric and sulfuric or phosphoric acids. As indicated in Fig. 11.5, there presently is no production, and little or no importation, of this type of fertilizer in the United States. The primary advantage of nitric phosphate processes is that no sulfur or less sulfur is required as compared with superphosphates or ammonium phosphates; this is particularly important during a shortage of sulfur, or in locations where sulfur must be shipped long distances. A variety of processes and equipment have been used in Europe since the late 1930s.[17] Also there are a number of plants in Central and South America and in Asia. In past years, there have been several U.S. ventures into production, but none has competed successfully with phosphoric acid–based processes. Production of nitric phosphates is complex. Simple substitution of nitric acid in a superphosphate-type acid–rock reaction is not feasible because (1) decomposition of the nitric acid would occur and cause noxious fumes and loss of nitrogen, and (2) the product would be extremely hygroscopic and unstable.

One process modification that overcomes these difficulties uses only nitric acid for acidulation; the extraction slurry is cooled to crystallize calcium nitrate, which then is removed by centrifugation. This process is referred to as the Odda process. The calcium nitrate by-product is either sold as a fertilizer or converted to coproduct ammonium nitrate. The extraction slurry, with its lowered calcium content, then is ammoniated and granulated. Various process modifications utilize a rotary drum, pug mill, or spray drum (Spherodizer) for granulation of the slurry. NPK nitric phosphate grades also have been produced by a prilling process in a few European plants. In earlier years, a main disadvantage of nitric phosphate processes was the low water solubility of phosphate in the products, as the main phosphatic constituent was dicalcium phosphate. The use of supplemental phosphoric acid, or "deep cooling" by refrigeration to remove a higher proportion of calcium nitrate in the Odda-type processes, now allows water solubility of 60 percent or higher. The remaining 40 percent of the phosphate,

although water-insoluble, is almost completely citrate-soluble, and this is "available" as plant nutrient. Popular grades of nitric phosphates include 14-14-14, 22-11-11, 20-20-0, and 16-13-0.

Diammonium Phosphate (DAP). Ammonium phosphates did not come on the fertilizer scene in significant quantities until the early 1960s; however, they have rapidly become the leading form of phosphate fertilizer in the United States and in the world. Almost all new phosphate fertilizer complexes built in recent years, and those planned, are for the production of ammonium phosphate as the major product. In 1960, U.S. production of ammonium phosphates was equivalent to only 270,000 tons of P_2O_5. By 1974, production was up to 2.1 million tons of P_2O_5 per year, and by 1990 it has reached 6.4 million tons/year. Ammonium phosphates now provide two-thirds of total U.S. phosphate fertilization.

Prior to the late 1940s, diammonium phosphate $(NH_4)_2HPO_4$ was considered to be too unstable for use as a fertilizer. However, at about that time, exploratory work by TVA indicated otherwise. The first demonstration-scale diammonium phosphate made by TVA for fertilizer use was a monocrystal product made by ammoniation of pure electric-furnace acid in a vacuum crystallizer.[18] Demonstration programs showed this material to be stable and to be an excellent high-analysis fertilizer. Small amounts of similar product were produced by others using electric-furnace acid and by-product ammonia. However, general acceptance of diammonium phosphate by the fertilizer industry was not possible until procedures could be developed to produce it from wet-process phosphoric acid instead of the relatively expensive electric-furnace acid. This occurred in 1960–61 with the development and patenting of the "TVA process" for production of granular 18-46-0 grade DAP from wet-process acid.[19,20] This relatively simple process was rapidly adopted by the industry, and remains the basic process for production of this major phosphate fertilizer. Many granular DAP plants have production capacities of about 50 tons/hr, with some going as high as 70 to 100 tons/hr.

The flow diagram of a typical granular DAP production unit of the TVA type that has become standard in the industry is shown in Fig. 11.11. Wet-process phosphoric acid of

Fig. 11.11. TVA process for production of granular diammonium phosphate. (*Courtesy TVA.*)

about 40 percent P_2O_5 content (often a mixture of 54% P_2O_5 acid and acid from the scrubbing circuit of 28–30% P_2O_5 content) is fed to a preneutralizer vessel. Anhydrous ammonia is sparged into this preneutralizer through open-end pipes that project through the walls. A key feature of the process is that this neutralization of the acid is carried to a closely controlled NH_3:H_3PO_4 mole ratio of about 1.4. Ammoniation to this degree provides maximum solubility of ammonium phosphate, as is shown by the solubility curve of Fig. 11.12. The heat of the ammoniation reaction evaporates considerable water, with the result that the water content of the slurry is reduced to only 16 to 20 percent. The slurry still is fluid and pumpable, however, because of the high solubility at a 1.4 mole ratio.

The preneutralizer slurry is pumped at a controlled rate to a rotary-drum, TVA-type ammoniator-granulator; there it is distributed onto a rolling bed of recycled solids. The most commonly used metering system for the preneutralized slurry is a variable-speed centrifugal pump with automatic control from a magnetic flowmeter. Some plants have had success with a magnetic flowmeter and automatic control valve of a special ball type.

Ammonia is sparged beneath the bed in the rotary drum to ammoniate the slurry further to near the theoretical DAP ratio of 2.0; the usual finishing NH_3:H_3PO_4 mole ratio is 1.85

Fig. 11.12. Effect of NH_3:H_3PO_4 mole ratio on solubility in the system NH_3–H_3PO_4–H_2O.

to 1.94. Metallic phosphate impurities such as iron, aluminium, and calcium phosphates account for most of the deviation from theoretical. This ammoniation in the granulator drastically decreases ammonium phosphate solubility (Fig. 11.12), and thus promotes solidification and granulation. Ammonia absorption is not 100 percent complete in either the preneutralizer or the granulator. However, the unabsorbed ammonia is recovered in a scrubbing circuit containing phosphoric acid of about 30 percent P_2O_5 content. Solid recycle to the granulator drum consists of undersize from screening of dried product plus some crushed, dry product and oversize. These combined materials are recycled at a rate of 5 to 7 lb per pound of product, and variation of this recycle rate is the primary method of controlling granulation efficiency.

Product discharged from the granulator is dried with moderate heat to a 180 to 190°F product temperature. Higher drying temperatures must be avoided because of a relatively low melting point for DAP. Most plants screen hot and cool only the product fraction because the material is not too sticky for recycle without cooling. Rotary coolers or fluidized-bed–type coolers are used. The product, with a moisture content of 1.5 to 2 percent, does not require a conditioner. It has excellent storage and handling properties in bags or in bulk. It has been shown that the iron and aluminium phosphate impurities in the product are present in a gel-like structure that favorably hardens the granules and increases their resistance to moisture.[21] The critical relative humidity of diammonium phosphate made from wet-process acid is about 70 percent; thus it is relatively nonhygroscopic.

Construction materials other than mild steel are required only for the acid lines, the preneutralizer, the slurry handling system, and the scrubbing circuits. Type 316L stainless steel or rubber- and brick-lined mild steel is used for the preneutralizer. Type 316L stainless is used also, for the slurry pumps and piping. Fiberglass-reinforced polyester plastic and high-density polyvinyl chloride sometimes are used for wet-process acid pipes and for

scrubbers. Scrubbers also may be constructed of rubber-lined mild steel.

Monoammonium Phosphate (MAP). "Diammonium Phosphate," "DAP," and the grade designation "18-46-0" have become extremely familiar terms in the world fertilizer industry. However, substantial interest has developed also in fertilizer use of monoammonium phosphate, $NH_4H_2PO_4$. Agronomically, MAP is favored where soils are mainly alkaline, for example, in Canada and Pakistan. From an economic standpoint, the relatively high P_2O_5 contents of typical MAP grades (11-52-0 to 10-54-0) make them attractive where the primary interest is in producing and shipping phosphate; the MAPs provide a higher P_2O_5 payload than does DAP. In the year ending June 30, 1990, granular monoammonium phosphates applied to the soil directly or in dry blends accounted for 17 percent of total U.S. fertilizer P_2O_5 usage. Smaller, additional amounts of MAP were used also, in formulating some granular mixed fertilizers and fluid mixtures.

TVA developed two comparatively minor modifications of the granular DAP process to allow production of granular MAP.[22] In one method, the acid is ammoniated to an NH_3:H_3PO_4 mole ratio of only about 0.6 in the preneutralizer and then to about 1.0 in the granulator drum. The 0.6 ratio is, like the 1.4 used in DAP production, a high-solubility point in the ammonium phosphate system (Fig. 11.12); the subsequent ammoniation to 1.0 in the granulator decreases solubility and thus promotes granulation. In the other procedure, which has been preferred by the industry, acid in the preneutralizer is ammoniated to the high-solubility NH_3:H_3PO_4 ratio of about 1.4, as in DAP production; then, additional wet-process acid is distributed onto the bed in the granulator to adjust back to the low-solubility MAP mole ratio of about 1.0. With either modification, the remainder of the process is the same as for DAP, except that, because of the higher temperature stability of MAP, a higher drying temperature can be used to increase the production rate for MAP.

Starting about 1968, simple processes were developed for production of nongranular (sometimes called powdered) MAP. Chief developers were Fisons and Scottish Agricultural Industries (SAI) in the United Kingdom, Swift in the United States, and Nissan in Japan. These processes involve variations in the simple reaction of wet-process acid with ammonia followed by spray-type drying with air. Based on this work, a number of comparatively low-cost units for production of nongranular MAP have been built commercially, including plants in the United Kingdom, the Netherlands, Japan, Australia, Spain, United States, Brazil, and Iran. This intermediate usually is shipped to other plants, where it then is used in the production of NPK solid or fluid fertilizers. Thus far, however, the nongranular MAP has not attained the popularity that was predicted for it in the mid-1970s.

POTASSIUM SALTS*

The element potassium has been recognized to be beneficial to plant growth since J. R. Glauber in the Netherlands first proposed, in the middle of the seventeenth century, that saltpeter (KNO_3) was the "principle" of vegetation. This salt, derived from the leaching of corral soils, was thus the first chemical compound intentionally applied to crops to satisfy their nutrient needs. Justus von Leibig had established the essentiality of the element for plant growth in his pioneer work published in 1840. The potassium or potash chemical industry dates from the year 1861, following the discovery by German chemists of a process for recovery of KCl (muriate of potash) from rubbish salts, the residues remaining from the extraction of sodium chloride (NaCl). World consumption of K now exceeds 28 million metric tons of K_2O equivalent annually. About 93 percent of this is consumed in fertilizers, the balance going to process use in industry.

*Sections on potassium were written by Dr. John A. Stewart, Libertyville, IL.

The functions of potassium in the plant are manifold. This element serves to activate or catalyze a host of enzyme actions, to facilitate the transport of nutrients and assimilates in the xylem and phloem, to maintain the structural integrity of the plant cell, to regulate turgor pressure, to mediate the fixation of nitrogen in leguminous plant species, and to protect plants to some degree from certain plant diseases. Plants affected by a deficiency of potassium usually are stunted and spindly, and exhibit a rather distinctive chlorosis and necrosis that is most prominent in the lower leaves (see Fig. 11.13). Crop yields may be severely reduced.

Potassium is a relatively abundant element in the earth's crust, ranking seventh in concentration. It is widely distributed geographically and is commonly found in association with sodium compounds. The feldspars, muscovite (white mica), granite, and gneiss are rich sources; but because they are siliceous and refractory, they are difficult and costly to convert to forms suitable for use as fertilizers. The common potassium minerals are listed in Table 11.3.

In the trade, potassium-containing compounds or salts that are applied primarily as a source of this nutrient or that find use in industry are referred to as potash. The name derives from an early production method in which potassium carbonate, leached from wood ashes, was crystallized by evaporating the leachate in large iron pots. The salt potassium chloride (muriate, muriate of potash, or KCl) is now the major source of the element (95%); other important salts are potassium sulfate (potassium sulphate, K_2SO_4, or sulphate of potash), potassium magnesium sulfates of varying K/Mg ratios, and potassium nitrate (KNO_3). The fertilizer industry expresses the potassium content of fertilizer salts in terms of the potassium oxide (K_2O) content, not as the K content. Muriate thus contains 60 percent+ of K_2O—which equates with 49.8 percent+ of potassium (K).

The potash industry is based on very large deposits of water-soluble potassium minerals resulting from the evaporation of shallow seas or natural brine lakes over a geological time span. These evaporites normally are located at depth in the earth's mantle. Ironically, the best deposits are found in areas quite remote from the more productive agricultural regions, which thus are the areas most in need of this element. Typically the deposits are to be found in horizontal tabular bodies or beds and may occur at depths up to 2100 meters or more. The beds may be only a few centimeters to a few meters thick, but commercial production is limited to strata that are at least one meter thick. As with hard-rock mining, the potash-bearing ores are extracted or harvested with continuous mining machines and brought to the surface through vertical shafts. Where the ores occur below a depth of about 1100 meters, or where the beds exhibit geological anomalies (e.g., folding), the potash is dissolved in a brine solution and pumped to the surface for recovery using solution mining techniques.

Potassium Minerals

The ore zone or stratum typically contains potassium or potassium–magnesium minerals along with halite (sodium chloride). Muriate of potash (M-O-P in the trade) is refined from sylvinite ore, a mechanical mixture of potassium chloride (KCl) and sodium chloride (NaCl). Because the latter salt is injurious to most crop plants, the KCl (sylvite) must be separated from the NaCl (halite).

For coarse-grained ores, physical methods employing froth flotation and/or heavy-media methods are employed to achieve separation. The heavy medium may consist of, for example, pulverized magnetite in a brine solution. The specific gravity of the medium is adjusted so that it falls between that of KCl (1.99 h/cm^3) and that of NaCl (2.17 g/cm^3).

When the crushed ore is placed in this medium, the K values are floated off, and the contaminating Na values sink and are drawn off and rejected from the bottom of the flotation vessel. Entrapped magnetite from both the product and waste streams is easily recovered magnetically and returned to the flotation vessel.

(A)

(B)

Fig. 11.13. Photographs of corn (A) and soybean (B), showing symptoms of potash deficiency.

TABLE 11.3 Common K Minerals (Zandon)

Mineral	*Formula*	*K Content*	*K_2O Content*
		g kg^{-1}	
Sylvite	KCl	524.4	631.7
Carnalite	$KCl \cdot MgCl_2 \cdot 6H_2O$	140.7	169.5
Kainite	$KCl \cdot MgSO_4 \cdot 3H_2O$	157.1	189.2
Langbeinite	$K_2SO_4 \cdot 2MgSO_4$	188.5	227.0
Leonite	$K_2SO_4 \cdot MgSO_4 \cdot 4H_2O$	213.3	256.9
Schoenite	$K_2SO_4 \cdot MgSO_4 \cdot 6H_2O$	194.2	233.9
Polyhalite	$K_2SO_4 \cdot MgSO_4 \cdot 2CaSO_4 \cdot 2H_2O$	129.7	156.2

Source: Stewart, J. A., in *Potassium in Agriculture*, R. D. Munson (Ed.), p. 83, ASA-CSSA-SSSA, 1985.

In froth flotation the crushed ore, which has been scrubbed to remove contaminating clay particles, first is treated with an aliphatic amine acetate and a froth-promoting alcohol. The water-repellent amine acetate is attracted preferentially to the sylvite particles. In a subsequent step the ore or pulp is maintained in suspension, and the coated sylvite particles are attracted to and entrained on the rising air bubbles generated by the flotation agitator. They rise to the surface and are collected by hydraulic and mechanical means. The K-depleted pulp then is discarded. The sylvite or muriate is collected, centrifuged to remove the hydraulic medium, dried, screened into various particle size ranges, and sent to storage. Figure 11.14 shows a modern potash mine/refinery operation.

Potassium–Magnesium Minerals

Other important sources of K for fertilizer use are the double salts kainite ($KCl \cdot MgSO_4 \cdot 3H_2O$) and langbeinite ($K_2SO \cdot 2MgSO_4$). The former is recovered from potash mines in Germany and Sicily; the latter is recovered and processed from mines in New Mexico (U.S.A.). These minerals supply important quantities of both magnesium and sulfur as well as potassium, and all are essential nutrients for plant growth. Kainite has a theoretical composition of 15.99 percent K (19.26% K_2O), 9.94 percent Mg, and 13.11 percent S, and langbeinite contains 18.85 percent K (22.7% K_2O), 11.71 percent Mg, and 23.18 percent S.

In the processing of langbeinite, the contaminating chlorides (NaCl, KCl) are removed by extracting the crushed ore with water. Centrifugation, drying, and screening follow to prepare the product in various size grades for the market. Where the mixed ore contains recoverable quantities of KCl as well as langbeinite, froth flotation and heavy-media methods may be employed to recover both K-containing salts.

Potassium Sulfate

Sulfate of potash (K_2SO_4 or SOP), unlike the earlier-discussed potash salts, does not occur as natural deposits. It can be recovered by fractional crystallization from the natural brines such as those of the Great Salt Lake in Utah, and Searles Lake in California. Here separation and recovery are achieved by solar evaporation in shallow ponds. These processes can be utilized only where a suitable brine source is available, and where solar evaporation rates are high.

A much more important source of K_2SO_4 is the product resulting from either a controlled decomposition reaction of a complex sulfate salt, reaction of the salt with KCl, or both. Thus kainite ($KCl \cdot MgSO_4 \cdot 3H_2O$) may be reacted with water in a two-step reaction, first to form schoenite ($K_2SO_4 \cdot MgSO_4 \cdot 6H_2O$), with the schoenite then reacted with KCl to form K_2SO_4.

Langbeinite likewise is reacted with KCl to form sulfate of potash, according to the

Fig. 11.14. Aerial view of IMC Fertilizers, Inc. potash mine/refinery at Esterhazy, Saskatchewan. Included in the photo are the head frame, cone storage buildings, and rail shipping facilities. The mine workings are 955 meters beneath the surface. (*Courtesy IMC Fertilizers, Inc.*)

following reaction:

$$\underset{\text{langbeinite}}{K_2SO_4 \cdot 2MgSO_4} + \underset{\text{potassium chloride}}{4KCl} \rightarrow \underset{\text{potassium sulfate}}{3K_2SO_4} + \underset{\text{magnesium chloride}}{2MgCl_2\,(\text{solution})}$$

Approximately 50 percent of world K_2SO_4 production is derived from the reaction of KCl with either kainite or langbeinite.

For those producers who do not have access to such suitable sulfate salts to serve as the sulfate donor, the so-called Mannheim process, in which sulfuric acid is reacted with KCl, may be utilized. The following reaction unfortunately generates the corrosive substance hydrochloric acid as a coproduct:

$$\underset{\text{potassium chloride}}{2KCl} + \underset{\text{sulphuric acid}}{H_2SO_4} \rightarrow \underset{\text{potassium sulfate}}{K_2SO_4} + \underset{\text{hydrochloric acid}}{2HCl}$$

Potassium sulfate, though relatively costly to produce, is a preferred source of K for some vegetable and fruit crops and tobacco—crops that are sensitive to the chloride forms.

Potassium Nitrate

This salt is produced in small amounts commercially in the United States, Israel, and

Norway by the reaction of KCl with nitric acid. Though the salt is an attractive fertilizer material, containing the two essential nutrients nitrogen and potassium in plant-available form, it is expensive to produce and thus enjoys only a limited market. Small amounts of a crude potassium nitrate also are marketed from natural deposits of caliche ($NaNO_3$) in Chile.

Potassium Phosphates

The production of such salts would seem to be a desirable objective because plant needs for both potassium and phosphate would be satisfied. Phosphoric acid can be reacted with KCl to produce the products, but in practice this route is not economical because energy use is high, reagents of high purity must be employed, K losses in the process can be severe, and the corrosive HCl coproduct demands the use of expensive corrosion-resistant equipment. Further it might fairly be asked whether such highly refined products can be justified for application to the soil. Such products intended for general agricultural use were produced in the United States some years ago, but their production now has virtually ceased.

MIXED FERTILIZERS

Preceding sections of this chapter have described the production of nitrogen fertilizers, phosphate fertilizers, and potash fertilizers from basic raw materials. Application of any of these fertilizers singly to the soil, without further processing, is referred to as "direct application." In the year ending June 30, 1990, only about 60 percent of the total fertilizer applied in the United States was applied directly; the remaining 40 percent was applied in the form of "mixed" fertilizers, that is, fertilizers that contain more than one of the major plant nutrients N, P, and K.[9] A breakdown by plant nutrient is as follows:

	% of Nutrient Applied by Indicated Method	
	Direct Application	*In Mixtures*[1]
Nitrogen	80	20
Phosphate	8	92[2]
Potash	65	35
Total nutrients	61	39

[1]Includes bulk blends.
[2]Includes ammonium phosphates applied alone.

These statistics show that for the phosphate fertilizers, application as mixtures (assuming the ammonium phosphates to be mixtures) far exceeds direct application, whereas for nitrogen fertilizers direct application far exceeds application in mixtures. For potash fertilizers, about 35 percent is applied as mixtures and 65 percent by direct application. The reasons for these patterns are both agronomic and economic. Agronomically, large applications of nitrogen alone—for example, in the "side dressing" of corn—often are of considerable importance. For phosphate, there usually is no similar agronomic need for single-nutrient application; so application in mixtures, with the resultant elimination of multiple handling, is preferred. Another factor, this one economic, that favors the direct application of nitrogen is the relatively low cost of anhydrous ammonia and nitrogen solutions, products that can be directly applied but ones whose use in mixtures is very limited. About 75 percent of the directly applied nitrogen is in the form of anhydrous ammonia, aqua ammonia, or nitrogen solution. The major solid nitrogen products (urea, ammonium nitrate, and ammonium sulfate), like the phosphate products, are used for the most part in mixed fertilizers, that is, in mixtures that contain two or more of the major plant nutrients (NPK).

Mixed fertilizers can be divided, by physical characteristics and production methods, into the following categories, all of which will be

discussed in some detail:

- Nongranular mixtures
- Compound granulars
- Bulk blends
- Fluid mixtures

Nongranular Mixtures

Until about 1950, essentially all the mixed fertilizers available to farmers were of the nongranular type. Production was by simple batch weighing and blending of available N, P, and K ingredients, usually in locations that were convenient with respect to rail or water receipt of raw materials and out-shipment of bagged mixtures to dealers in farm areas. The source of phosphate almost invariably was nongranular ordinary superphosphate (OSP) produced at the mixing location from in-shipped phosphate rock and locally produced sulfuric acid. In-shipment of elemental sulfur was required to produce the acid, usually by the now obsolescent "chamber" process. The major nitrogen ingredient used in the mixtures was fine-crystal by-product ammonium sulfate, and the potash source was nongranular potassium chloride. In most plants, the superphosphate was "ammoniated" by reaction with ammonia solution, either in a separate operation before mixing with the other ingredients or during the mixing operation. Such ammoniation was beneficial in reducing the corrosiveness (acidity) of the superphosphate and in improving its physical condition. The ammonia also provided nitrogen at relatively low cost, but the amount that could be absorbed by the superphosphate was limited. Organic materials often were included in the mixtures as nutrient sources, bulking agents, or physical conditioners. Materials used included fish meal, packinghouse wastes, dried blood, tobacco stems, cotton seed hulls, and others. Mineral materials such as kaolin clay, diatomaceous earth, and vermiculite sometimes were included as conditioners. Batch-type drum mixers of 1- to 3-ton capacity were used, with the ammonia solution piped directly into the mixer. Mixture grades were low by present-day standards, and storage and handling properties of these nongranular mixtures were generally poor. Caking problems, in particular, were considerable. Production of this type of fertilizer now persists in only a few locations. The quantity produced in the United States is very small.

Compound Granulars

TVA Continuous Ammoniator-Granulator Process. Following the close of World War II in 1945, farmer demands developed for much increased quantities of fertilizer and for products with better handling properties. Granulation, which provides an increase in particle size and the elimination of fines, was a promising avenue for the physical improvement that was pursued by researchers in both the United States and Europe. A milestone in this movement to granulation was the development, patenting, and demonstration by TVA, in 1954, of a continuous rotary-drum granulation process for mixed fertilizers.[23,24]

A flow diagram of a plant employing this process is shown in Fig. 11.15. The rotary drum granulator in this process is essentially the same as was pictured for the diammonium phosphate granulation process. In early versions of the process, feed to the drum consisted of (1) crushed oversize product and fines, (2) nongranular ordinary superphosphate, (3) nongranular potassium chloride, (4) ammonium sulfate crystals, (5) sulfuric acid, and (6) anhydrous ammonia or an ammonia-containing nitrogen solution. The acid and the ammonia or ammoniating solution were sparged under the rolling bed of solids in the drum with a sparger arrangement, as shown in Fig. 11.16. With the proper balance of liquid and solid feeds, granulation occurs as a result of the rolling action in the drum, and it is completed, with hardening, in a rotary drum dryer.

This process was rapidly and widely adopted by mixed-fertilizer producers throughout the United States. By the mid-1960s the number of U.S. plants employing the process had

Fig. 11.15. TVA-type ammoniation–granulation plant for NPK mixed fertilizers. (Fertilizer Manual, *International Fertilizer Development Center, Muscle Shoals, AL, Courtesy TVA.*)

Fig. 11.16. Sparger location under rolling bed in TVA-type ammoniator–granulator. (Fertilizer Manual, *International Fertilizer Development Center, Muscle Shoals, AL, Courtesy TVA.*)

reached 200, with an average production rate of 100,000 tons/year per plant. The process was particularly attractive because: (1) the raw materials were essentially the same as materials that were already available at most batch-type nongranular plants, (2) higher proportions of relatively cheap anhydrous ammonia or nitrogen solution could be used, (3) continuous-type processing allowed higher production rates, and (4) product physical properties were much improved as a result of granulation. Today, the same basic process, but with numerous modifications, is used in most U.S. mixed-fertilizer granulation plants. However, the number of plants has dwindled to about 20, in favor of bulk blending and fluid fertilizers. One early modification to the process was the addition of phosphoric acid as a feed material, either along with or in place of sulfuric acid. Other feed materials used in some variations include triple superphosphate

(TSP), monoammonium phosphate, ammonium nitrate, and urea.

Use of a Preneutralizer (Slurry Granulation). A major improvement to the basic TVA granulation process was the incorporation in most plants of a preneutralizer vessel for prereaction of phosphoric acid with ammonia. This procedure, as applied to the production of granular diammonium phosphate (DAP), was described earlier and was pictured in Fig. 11.11. As in the DAP process, the preneutralizer used in the production of mixed fertilizers is operated at an NH_3:H_3PO_4 mole ratio of about 1.4 to give maximum ammonium phosphate solubility and thus maximum fluidity of the slurry pumped to the granulator. Subsequent further addition of ammonia in the granulator sharply decreases solubility, and thus promotes solidification and hardening of granules. As in ammonium phosphate production, the iron and aluminium impurities normally present in wet-process phosphoric acid contribute significantly to the hardening of granules during ammoniation.[21] Sulfuric acid, when used, is fed directly to the granulation drum, not to the preneutralizer. The use of the preneutralizer allows inclusion of higher proportions of acid in formulations. Also, better use is made of the heat of the ammonia–acid reaction for the evaporation of water, with resultant savings in dryer fuel.

Use of a Pipe Reactor (Melt Granulation). A more recent development that has been adopted in most U.S. granulation plants involves the substitution of a pipe reactor or a pipe-cross reactor for the preneutralizer vessel as the locale for the ammonia–acid reaction.[25–29] A simple pipe reactor differs from the pipe-cross reactor in that only one acid inlet port, usually for phosphoric acid, is provided.

The salient feature of the pipe reactor and pipe-cross reactor systems is that the heat of the ammonia–acid reaction is confined to the pipe and is efficiently utilized there to vaporize essentially all the moisture in the feed acid. The resultant steam exits the pipe discharge end, in the granulator, and is swept away in an air stream. The ammonium phosphate melt, which is of very low moisture content, also exits the pipe and is discharged onto the rolling bed of dry solids in the granulator. This melt serves as the binder for granule formation. The reaction temperature in the pipe reactor usually is held below 300°F to avoid a buildup of troublesome scale in the pipe. In view of this temperature limitation, there is little or no conversion of the phosphorus to polyphosphate forms. By allowing the pipe temperature to rise to 400°F or higher, polyphosphate can be formed without scaling, but this mode of operation seldom is used in mixed-fertilizer granulation. TVA has used such high pipe temperatures to produce 11-57-0 and 28-28-0 grade ammonium polyphosphate granules, in which 15 to 25 percent of the P_2O_5 is in polyphosphate form.[25] The physical properties of these products are exceptionally good, and the polyphosphate content makes them particularly suitable for dissolution to form fluid fertilizers. Because of the very low moisture content of the pipe reactor melts, even at the lower operating temperatures, the normally expensive drying of finished granular product usually is not needed; passage of the product through a cooler with an air sweep is sufficient for final drying. This results in a major fuel saving, which is a very significant economic advantage.

Pipe and pipe-cross reactors are being adopted by many plants; as of May 1990, about 15 were known to be in use in U.S. mixed-fertilizer granulation plants. Adaptability of the pipe reactor to DAP production is being explored, but is not yet firmly established.

Steam Granulation. A simple method for granulating dry mixtures of fertilizer materials was developed in Europe in the early 1950s and still is employed there to some extent. In this procedure, which is pictured in Fig. 11.17, the feed materials all are finely ground to ensure their incorporation in the granules. A dry mixture is made and then is moistened in a rotary drum or pan-type granulator with

Fig. 11.17. Steam granulation process with feed of dry, pulverized fertilizer materials. (Fertilizer Manual, *International Fertilizer Development Center, Muscle Shoals, AL.* Courtesy TVA.)

either water spray, steam, or both. This added moisture forms a salt solution by partial dissolution of the mixture ingredients. The granules form because of the plasticity and rolling action in the granulator. Initially, these granules are very soft; but, with further rolling and drying in a subsequent drying operation, salt bridging imparts the required granule hardness. In contrast to the TVA ammoniation–granulation process, the steam granulation process usually causes no significant reaction between mixture ingredients. In Europe, this feature is of considerable significance when superphosphate is used as a feed material. Ammoniation of superphosphate, which occurs in the TVA ammoniation–granulation process, converts part of the water-soluble P_2O_5 content of the superphosphate to water-insoluble forms that are not accepted as marketable phosphate fertilizer in some European countries. This conversion, however, is mainly to citrate-soluble forms, which are accepted as effective, marketable phosphate forms in the United States.

Bulk Blends

History and Growth. In the "compound" (mixed) fertilizer granulation processes just described, the starting materials are N, P, and K materials in the form of (1) nongranular solids, (2) liquids, or (3) gas (ammonia). When N, P, and K materials are all available from basic producers in granular form, as they now are, granular mixed fertilizers of almost any nutrient proportion (grade) can be made by simple proportioning and dry blending of the granular N, P, and K materials; further granulation or other processing is not required. This is the basis of the presently popular "bulk blending" system of mixed fertilizer preparation and distribution.

As can be seen from Figure 11.18, this system, since 1960, has taken a considerable share of the U.S. mixed-fertilizer market away from the compound granulation plants. As of 1990, bulk blending was furnishing about 62 percent of the mixed fertilizer used in the United States, as compared to about 15

Fig. 11.18. Changing patterns in the mixed-fertilizer market. (*Courtesy TVA.*)

percent as compound granulars and 23 percent as fluid mixtures.

The simplicity of the bulk blending procedure is illustrated in Fig. 11.19. The operations consist of (1) in-shipment of granular fertilizer materials, (2) temporary storage, (3) proportioning, (4) blending, and (5) movement of finished blend to the field for spreading.

Key factors in the rapid growth of bulk blending have been: (1) convenience in providing the farmer a wide choice of plant-food ratios or grades on a "prescription" basis in relatively simple, small mixing plants located in the immediate farm area; (2) economy due to handling in bulk rather than bags; and (3) provision of spreading and other farmer services by the local blender.

In this streamlined production and marketing system, the bulk blender is the retail dealer. The blender's service area usually is limited to a 30-mile radius around his or her plant, thus enabling a close relationship with local

Fig. 11.19. Bulk blending—a simple concept. (*Courtesy TVA.*)

farmers. The blender usually transports the fertilizer to the farms and applies it, for a fee. In addition to preparing and applying blends, bulk blenders also market and apply single-nutrient direct-application fertilizers, frequently including anhydrous ammonia and nitrogen solutions.

Typical blending plants produce and market only 2000 to 6000 tons of blend per year, but there are over 5000 plants located throughout farming areas. There also are a few plants with a much larger annual production, designed to serve relatively large areas. The capital investment in blending plants is comparatively low, but most blenders have sizeable investments in application equipment. Blending plants operate on a quite seasonal basis, and often the operators have supplemental business enterprises.

Intermediates Used in Bulk Blending. The bulk blending system came into being only after suitable N, P, and K intermediates became available in granular form. In particular, the introduction of granular diammonium phosphate (DAP) of 18-46-0 grade in about 1960, as described in a previous section, was a major catalyst to bulk blending. Characteristics of that product that make it an ideal material for shipping and subsequent blending are: (1) a high plant food content of 64 percent, (2) high phosphorus content of 46 percent, (3) excellent storage and handling properties, (4) compatability with essentially all other fertilizer materials, and (5) economy of production near phosphate deposits. Most NPK bulk blends are prepared with DAP as the phosphate source. Granular triple superphosphate (TSP) is used for no-nitrogen (PK) grades, but for NPK blends TSP has the disadvantage of being incompatible (reactive) with urea.

In the preparation of bulk blends, it is of great importance to ensure that all ingredients of the blend are reasonably matched in particle size.[30] The use of unmatched materials results in poor mixing, segregation during handling, and uneven distribution on the field. Nitrogen materials used with the greatest success in bulk blends are (1) granular urea, (2) granular ammonium nitrate, and (3) large-crystal or granulated ammonium sulfate. Prilled ammonium nitrate also can be made of large-enough particle size for good blending, but prilled urea normally is of a smaller size than desirable. The chief potash material used in blends is granular-grade potassium chloride, which is sized specially by potash producers for use in blends. Although granular-size potash usually consists of irregular crusher fragments rather than well-rounded granules, studies have shown that this particle shape difference does not promote segregation or otherwise impair blend quality.[31] A somewhat smaller-size, coarse-grade potassium chloride sometimes is used in blends because it is cheaper, but segregation problems are invited. Potassium sulfate and potassium magnesium sulfate are made in granular form for the blending of specialized formulations.

The addition of micronutrients to bulk blends poses special problems because of the small quantities usually added. Although many micronutrients are available in granular form, intergranule spacing on the field can be excessive when the granular micronutrient constitutes only a very small proportion of the blend. Procedures have been devised and used for adding finely powdered micronutrients to blends together with a small amount of liquid binder.[32] The result is that each blend granule is coated with micronutrient, and the field distribution thus is better.

Operating experience and TVA studies have identified several combinations of materials that are incompatible in blends and should be avoided: (1) unammoniated superphosphates react with urea, and the result is wetting and caking; and (2) urea in contact with ammonium nitrate forms an extremely soluble hygroscopic mixture that liquifies in the presence of only minute amounts of moisture. Therefore, these two materials can never be used together in the same blend. Essentially all other combinations of available fertilizer materials are compatible in blends.

Equipment and Operating Technique. There are many different arrangements of equipment in the thousands of blending plants throughout

the United States. The operations in such plants generally consist of: (1) receipt of the intermediate granular components in bulk by truck or railroad car, (2) storage of the intermediates separately in bins, (3) removal of the granular blend components from storage and weighing in desired proportions, (4) mixing to obtain uniformity, and (5) discharge of the mixed batch directly into a spreader truck or into a holding bin. The most popular type of mixer is the batch-type rotary drum.

Such mixing plants can be highly automated so that the output with a mixer of only one-ton capacity can be as much as 15 to 20 tons/hr. Mixers of several types are available with capacities up to 6 tons per batch.[33]

Problems in Bulk Blending. Bulk blending is a simple practice that can provide uniform mixtures in almost any desired proportions. However, unless proper materials are used and certain handling precautions are taken, segregation can occur, and nonuniform blends will result. In general, the blend-plant records of conformance to state analytical requirements have not been as good as those of compound granulation plants.[34] The factors involved and the precautions that are necessary have been described by TVA.[35] The main requirements are the use of granular materials with a well-matched range of particle size and the handling of the mixture after blending in ways that will minimize segregation. The mixing step is seldom a problem; it is quite easy to obtain a uniform blend by the use of proper materials and a few minutes of mixing. However, maintaining this uniformity during handling on the way to the farm and onto the soil requires precautions. Handling procedures that may cause segregation include coning (as occurs if the blend is allowed to drop from a discharge pipe onto a pile), vibration in hauling vehicles, and ballistic action imparted by some types of spreaders.

The problems of adding small quantities of micronutrient materials to bulk blends was mentioned earlier. This problem is being met by either using micronutrients in granular form or coating the blend particles with finely ground micronutrients and a liquid binder.[32]

Fluid Mixtures

Mention was made earlier of the significant quantity of nitrogen applied to farm lands in the form of urea–ammonium nitrate solution and aqua ammonia. Since the late 1950s, there also has been a growing practice of applying mixed fertilizers (NP and NPK) in fluid form. The growth curves in Fig. 11.18 show that by 1990 fluid mixtures had claimed about 23 percent of the mixed-fertilizer market. This growth paralleled the growth of bulk blending, at the expense of the compound-granular-type products. It is expected that, in the future, bulk blending and the production of fluids will continue their parallel growth patterns.

Numerous advantages have been cited[10] to explain the growth of fluid mixed fertilizers. Of these, the most significant is probably the ease and the precision with which fluids can be applied to the soil. When placement of the fertilizer is important, as in row fertilization, fluids often have an advantage. Also important is the adaptability of fluids to the homogeneous incorporation of micronutrients, herbicides, and insectiides. Another factor that encouraged the adoption of fluid mixtures was the prior existence of equipment and facilities for handling and applying nitrogen fluids; it was only logical for this practice to be expanded to include preparation and application of mixtures. There are some claims of agronomic superiority for fluid fertilizers over similarly applied solids, but such claims are debatable and are not widely accepted.

Solution Fertilizers. Fluid fertilizers in use today can be divided into two major categories: (1) solutions, in which the plant food content is entirely in solution; and (2) suspensions, in which the plant food content exceeds its solubility, and the excess is held in suspension, usually by addition of a suspending agent. The solution-type mixtures were the first to achieve commercial importance. Beginning about 1950, the practice of making

solution mixtures was begun, with pure electric-furnace orthophosphoric acid (54% P_2O_5) as the source of phosphorus. Ammoniation of this acid yielded an 8-24-0 grade ammonium orthophosphate solution that could be shipped to mixing plants in farm areas and used as a base for blending with other fertilizer materials. Blending usually was with urea–ammonium nitrate solutions (28–32% N) and a pure grade (solution grade) potassium chloride.

With the system that employed these raw materials, numerous advantages of fluid mixtures were established, but a need to improve the economics of the procedure was obvious. Pure electric-furnace acid is much more expensive than wet-process acid, so ways were sought to use the cheaper but impure wet-process acids. Direct ammoniation of wet-process acid causes precipitation of numerous impurities as an intolerable, gelatinous sludge. Also, plant-food solubilities are undesirably limited in the orthophosphate system, so the concentration (grade) of the product made with electric-furnace orthophosphoric acid was limited to 8-24-0. A series of developments by TVA led the way to solving both of these problems.

The problem of limited plant-food solubility was much alleviated by TVA's introduction of superphosphoric acid in 1957.[36] Superphosphoric acids are acids of high P_2O_5 content (up to 80% P_2O_5) that contain significant proportions of polyphosphates (non-ortho, polymerized phosphates). The data in Fig. 11.20 show the favorable effects of polyphosphate on plant-food solubility in ammoniated phosphoric acids. The superphosphoric acid first produced by TVA was a pure electric-furnace product of 76 percent P_2O_5 content, of which about 50 percent was in polyphosphate form. With ammoniation of

Fig. 11.20. Effect of polyphosphate level and $N:P_2O_5$ weight ratio on solubility of ammoniated phosphoric acids at 32°F. (*"Fluid Fertilizers,"* TVA Bulletin Y-185, *p. 76. Courtesy TVA.*)

this acid, it was possible to produce an 11-37-0 grade base solution instead of the 8-24-0 grade maximum made from ortho acid. Demonstration use of 11-37-0 base in a number of mixing plants showed that solution mixtures of relatively high concentration could be made. Also, it was found that the polyphosphate content enhanced the solubility of micronutrient additives by a sequestering action.

The unfavorable economics of the electric-furnace acid process, however, discouraged the commercial production of electric-furnace superphosphoric acid; so attention was directed toward concentration of the cheaper wet-process phosphoric acid to give wet-process superphosphoric acid. Satisfactory concentration methods were developed,[37] but the effects of acid impurities on handling properties of the concentrated acids limited the final P_2O_5 concentration and hence the polyphosphate content. The best handling properties were obtained with acids made from calcined rocks and with P_2O_5 concentration limited to 68 to 70 percent P_2O_5. In acid of this concentration, from 20 to 35 percent of the P_2O_5 is in the polyphosphate form, and the acid is referred to as low-conversion superphosphoric acid. Such acid today is the chief source of P_2O_5 for solution-type fertilizers.

The most popular scheme by which solution-type fertilizers are made from low-conversion wet-process superphosphoric acid is outlined in Fig. 11.21. The first step is ammoniation of the acid to produce ammonium polyphosphate base solution. By carrying out this ammoniation in a TVA-developed pipe reactor, a temperature of over 400°F is developed, which promotes an increase in polyphosphate content to 65 to 70 percent of total P_2O_5 in the ammoniated acid.[38,39] The grade of the resultant base is usually 10-34-0; an 11-37-0 grade can be made, but its storage properties are not as good, especially in cold weather. A sketch of a typical TVA pipe-reactor system for production of high-polyphosphate base solution from low-conversion wet-process superphosphoric acid is shown in Fig. 11.22.

More than 125 U.S. plants use this TVA pipe-reactor process to produce an estimated 2 million tons of 10-34-0 or 11-37-0 grade product per year. Other countries such as Belgium, France, and the Former Soviet Union (FSU) also use this process to produce ammonium polyphosphate solution. A considerable amount of the solution is used for direct application in the U.S. wheat belt and other areas in which potassium is not deficient. The rest is used in small mix plants to produce mixtures. The ammonium polyphosphate base solution is mixed with nitrogen solutions containing 28 to 32 percent N (urea–ammonium nitrate solutions) and sometimes also with potash to produce such clear liquid grades as 21-7-0, 7-21-7, and 8-8-8. As with bulk blending, the solution mixing operation usually is carried out in small mixing plants located in farm areas.

Suspension Fertilizers. Suspension fertilizers were developed chiefly as a means of breaking the grade barriers imposed on solution fertilizers by solubility limitations. The low solubility of potassium chloride was particularly troublesome, in that high-potash solution grades were not possible.

The feasibility of preparing and handling suspension-type fertilizers was demonstrated by TVA and others in 1959.[40,41] Two essential requirements for the preparation of a successful suspension were shown to be: (1) small particle size of the suspended solids, and (2) the addition of a gelling-type clay as a suspending agent along with sufficient high-shear agitation to properly disperse the clay. The most satisfactory gelling clay is attapulgite, obtained in south Georgia and north Florida.[42] Attapulgite is a hydrated magnesium aluminium silicate, which is composed of needle-shaped crystals so small that they can be seen only with use of high magnification. In the dry state, these crystals bond together in bundles as larger particles. These bundles must be disrupted and the individual crystals freed before they can assume the structural arrangement that is characteristic of a gel. When this disruption of the crystal bundles is carried out in the presence of an electrolyte, such as any of the common fertilizer salts present in fluid fertilizers, the individual

Fig. 11.21. Production methods for solution-type and suspension-type mixed fertilizers.

Fig. 11.22. TVA pipe reactor system for production of ammonium polyphosphate base solution from low-conversion superphosphoric acid. The product base solution is typically 10-34-0 with a 65–70% poly level. (*Courtesy TVA.*)

crystals immediately assume the regular structural orientation of a weak gel, and thus provide the desired suspending action.

Wetting of the clay by the fluid fertilizer causes some of the required disruption of crystal bundles and freeing of individual crystals, but not enough to provide good gelling without high-shear agitation. An addition of 2 to 3 percent of attapulgite is sufficient for most fertilizer suspensions. Instead of providing high-shear agitation for the entire batch of suspension, some operators carry out a pre-gelling operation in which high-shear agitation is applied only to the clay plus a limited amount of water or fertilizer solution. The pre-gelled mixture then is added to the major body of fertilizer with only mild mixing. Pre-gelled clay also is marketed as liquid clay and is transported in tank cars.

There is considerable variety in the types of fertilizer raw materials used commercially in preparation of suspension fertilizers. This is indicated in Fig. 11.21, where some of the more popular materials are listed. Likewise, production procedures and equipment vary considerably. Provision of a high-shear mixer is essential not only for gelling of the suspending clay but also for disintegration of

Fig. 11.23. Suspension fertilizer production using solid feeds (1–3% gelling clay also included in mix). (*Courtesy TVA.*)

any granular products such as diammonium phosphate or monoammonium phosphate that might be used in the suspension. A high-shear recirculation pump also is an aid to gelling of the clay.

Figure 11.23 shows the layout of a typical suspension mix plant in which all the major suspension ingredients are received as solids. This is presently a popular mode of operation. Raw material costs for such a plant often are less than for bulk blending because nongranular materials can be used. Piping for the optional addition of phosphoric acid and ammonia is shown. Such addition develops heat, which hastens the disintegration of some solids.

The storage and handling properties of suspensions can be enhanced by inclusion of some polyphosphate in the suspension. Low-temperature storage properties, in particular, are enhanced. One popular method of providing polyphosphate is to supply all or part of the P_2O_5 as 10-34-0 ammonium polyphosphate base solution, such as is used in solution fertilizers. A second method, shown as an option in Fig. 11.21, is the use of a pipe reactor to prepare 9-32-0 or 12-36-0 grade ammonium polyphosphate base suspension from relatively inexpensive wet-process orthophosphoric acid (54% P_2O_5), ammonia, and gelling clay. This process was introduced by TVA in 1982;[43] a flow diagram is shown in Fig. 11.24. About 20 percent of the P_2O_5 in the 9-32-0 or 12-36-0 product is in polyphosphate form. The clay content is 2 percent. When these base solutions are used for mixing with other suspension ingredients, their clay content often is sufficient to provide all the clay required in the final suspension. Commercial acceptance of 9-32-0 or 12-36-0 base suspension preparation has not yet occurred. At present, it apparently is most economical to furnish polyphosphate as 10-34-0 ammonium polyphosphate solution.

Fig. 11.24. TVA pipe-reactor process for production of 9-32-0 grade ammonium polyphosphate base suspension from wet-process orthophosphoric acid. (*"Fluid Fertilizers,"* TVA Bulletin Y-185, *p. 90. Courtesy TVA.*)

REFERENCES

1. Free, W. J., Bond, B. J., and Nevins, J. L., "Changing Patterns in Agriculture and Their Effect on Fertilizer Use," TVA Fertilizer Conference, July 27–28, 1976, *TVA Bulletin Y-106*, Tennessee Valley Authority, Muscle Shoals, AL.
2. *Farm Chemicals Handbook*, p. 134, Meister Publishing Co., Willoughby, OH, 1970.
3. "Summary of State Fertilizer Laws," Product Quality and Technology Committee, The Fertilizer Institute, Washington, DC (1991).
4. *Official Methods of Analysis*, Association of Official Analytical Chemists, Inc., Arlington, VA, 1990.
5. "Micronutrients in Agriculture," Soil Science Society of America, Inc., Madison, WI (1972).
6. Hignett, T. P., "History of World Fertilisers and Manufacturing Processes," *Indian Chemical Manufacturer*, **9**, (4), 13–17 (1971).
7. *The Fertilizer Handbook*, The Fertilizer Institute, Washington, DC, 1982.
8. McVicar, M. H., *Using Commercial Fertilizers*, 3rd ed., The Interstate Printers and Publishers, Inc., 1970.
9. Hargett, N. L., and Berry, J. T., "Commercial Fertilizers," *1990 TVA/NFERC Bulletin Y-216*, Tennessee Valley Authority, Muscle Shoals, AL (1990).
10. Hignett, T. P. (Ed.), *Fertilizer Manual*, International Fertilizer Development Center, Muscle Shoals, AL, 1978.
11. "Storage and Handling of Anhydrous Ammonia," *Bulletin K61.1*, American National Standards Institute, Philadelphia, PA.
12. International Fertilizer Development Center (IFDC), "Compaction–Alternative Approach for Granular Fertilizers," *IFDC Technical Bulletin T-25* (1983).
13. U.S. Patent 4,589,904.
14. Lehr, J. R., and McClellan, G. H., "Phosphate Rocks; Important Factors in Their Economic and Technical Evaluation," *CENTO Symposium on the Mining and Beneficiation of Fertilizer Minerals*, pp. 192–242, Nov. 19–24, 1973.
15. Young, R. D., and Achorn, F. P., "Trends in U.S. Fertilizer Technology," *TVA Bulletin Y-133*, pp. 5–7, Tennessee Valley Authority, Muscle Shoals, AL (Aug. 1978).
16. "The Role of Phosphorus in Agriculture," *Am. Soc. Agron., Crop Science Soc. Am., and Soil Science Soc. Am.*, pp. 198–199, Madison, WI, 1980.
17. Hignett, T. P., "Nitrophosphate Processes Advantages and Disadvantages," *Proc. 15th Annual Meeting Fert. Ind. Round Table*, pp. 92–95, 1965.

18. Getsinger, J. G., Houston, E. C., and Achorn, F. P., "Production of Diammonium Phosphate by Continuous Vacuum Crystallization," *J. Agr. Food Chem.*, **5**, 433–436 (June 1957).
19. Young, R. D., Hicks, G. C., and Davis, C. H. "TVA Process for Production of Granular Diammonium Phosphate," *J. Agr. Food Chem.*, **10**, 442–447 (Nov. 1962).
20. U.S. Patent 3,153,574.
21. "New Developments in Fertilizer Technology," 8th Demonstration, *TVA Bulletin Y-12*, Tennessee Valley Authority (1970).
22. Young, R. D., and Hicks, G. C., "Production of Monoammonium Phosphate in a TVA-Type Ammonium Phosphate Granulation System," *Comm. Fert.*, **114**, (2), 26–27 (Feb. 1967).
23. Yates, L. D., Nielson, F. T., and Hicks, G. C., "TVA Continuous Ammoniator for Superphosphates and Fertilizer Mixtures," Part I and Part II of *Farm Chemicals*, (Aug. 1954).
24. U.S. Patent 2,741,545.
25. "New Developments in Fertilizer Technology," 9th Demonstration, *TVA Bulletin Y-50*, Tennessee Valley Authority (1972).
26. "New Developments in Fertilizer Technology," 11th Demonstration, *TVA Bulletin Y-107*, Tennessee Valley Authority (1976).
27. Medbery, J., "Experience with Pipe-Cross Reactor," *Proc. Fert. Ind. Round Table*, pp. 52–55 (1977).
28. Baggett, K. J., and Brunner, D. J., "MFA's Experience with TVA Cross-Pipe Reactor," *Proc. Fert. Ind. Round Table*, pp. 64–70 (1977).
29. Fittell, R. S., Hollingworth, L. A., and Forney, J. G., "Manufacture of Granular Ammonium Phosphates Using a Pipe Reactor Process," *Proc. Fert. Ind. Round Table*, pp. 70–81 (1977).
30. Hoffmeister, G., Watkins, S. C., and Silverberg, J., "Bulk Blending of Fertilizer Material: Effect of Size, Shape and Density on Segregation," *J. Agr. Food Chem.*, **12**, 64–69 (Jan./Feb. 1964).
31. Hoffmeister, G., "Compatibility of Raw Materials in Blended Fertilizers—Segregation of Raw Materials," *Proc. 12th Ann. Meeting Fert. Ind. Round Table*, pp. 83–88, 1962.
32. Young, R. D., "Providing Micronutrients in Bulk-Blended, Granular, Liquid and Suspension Fertilizer," *Comm. Fert.*, **118**, (1), 21–24 (Jan. 1969).
33. Young, R. D., "Production of Compound Fertilizers from Intermediates in Local Plants," *TVA Bulletin Z-30*, pp. 12–20, Tennessee Valley Authority (1971).
34. "Results of Fertilizer Quality Control Survey—1989," Assn. of Am. Plant Food Control Officials and the Fertilizer Institute, Washington, DC (1989).
35. "Quality Control in a Bulk Blending Plant," TVA Circular Z-49, Tennessee Valley Authority (1974).
36. McKnight, D., and Striplin, M.M., "Phosphoric Acid of High Concentration," *J. Agr. Chem.*, **13**, 33–34 (Aug. 1958).
37. Phillips, A. B., "Concentration Wet-processes Acid to Superphosphoric Acid," *Farm Chem.*, **126**, 36, 38, 62 (June 1963).
38. Meline, R. S., Lee, R. G., and Scott, W. C., "Use of a Pipe Reactor in Production of Liquid Fertilizers with Very High Polyphosphate Content," *Fert. Soln.*, **16**, (2), 32–45 (Mar./Apr. 1972).
39. U.S. Patent 3,775,534.
40. Walters, H. K., Jr., "Salt Suspension Fertilizers," *Comm. Fert.*, pp. 25–26 (Sept. 1959).
41. Newsom, W. S., Jr., "Suspension Fertilizers," *Solutions*, pp. 30–35 (Jan.–Feb. 1960).
42. Sawyer, E. W., Polon, J. A., and Smith, H. A., "The Stabilization of Suspension Fertilizers with Colloidal Attapulgite," *Solutions*, pp. 36–43 (Jan.–Feb. 1960).
43. Mann, C., II, McGill, K. E., and Jones, T. M., *I&EC Product Research and Development*, pp. 488–495 (Sept. 1982).

12

Salt, Chlor-Alkali, and Related Heavy Chemicals

Indresh Mathur*

SODIUM CHLORIDE

Common salt, sodium chloride, occurs in nature in almost unlimited quantities. It is a direct source of chlorine, caustic soda (sodium hydroxide), sodium chlorate, synthetic soda ash (sodium carbonate), sodium metal, and sodium sulfate. Indirectly, it is also the source of hydrochloric acid and a host of sodium salts. It has an imposing list of uses, placing it among the more important substances in the economic world. It is used to preserve meat, fish, and hides; it is a necessary component of the animal diet; it is used in refrigeration systems; and large quantities are used for ice control on highways in colder climates. Salt is used by the soap maker to separate soap from glycerine and lye, and by the dye manufacturer to precipitate dye products. In addition, salt is used extensively for the regeneration of water-softening resins (see Table 12.1).

Salt is mined as a solid in shaft mines with depths of 500 to 2000 feet, in Michigan, Ohio, New York, Kansas, Louisiana, and Texas. Run-of-the-mine salt contains 98 to 99 percent NaCl. Over 25 percent of the world's salt is produced in the United States.

More frequently, salt is obtained by solution mining. Typically, this is achieved by drilling a well into a salt formation, installing concentric piping into the well, pumping water in one pipe, and retrieving nearly saturated brine from the other pipe. The brine concentration is controlled by the rate of pumping and is kept slightly undersaturated to avoid salting up the brine lines at the wellhead. The brine is purified by chemical treatment, settling, and filtration. Such artificial brines permit a cheaper operating cost and are well adapted to the manufacture of synthetic soda ash, chlor-alkali products, and table salt.

Very extensive production of artificial brines is carried out along the U.S. Gulf Coast, principally in Texas and Louisiana, where huge domelike deposits of rock salt exist in readily accessible areas. Some of these domes are 4 miles in diameter and over 8 miles deep,[1] containing over 100 billion tons of salt each. In order to mine salt efficiently and safely and to utilize the subsequent cavern for storage,

*Dow Chemical Canada Inc., Sarnia, Ontario, Canada

TABLE 12.1 1986 Uses of Salt (Millions of Short Tons) and 1992 Projected Demand[3]

	1986	*1992*
Chlor-alkali	19.0	10.0–20.5
Highway ice control	10.5	10.0–11.0
Food processing	1.2	1.0–1.5
Feed and mixes	1.3	1.4
Other chemicals	0.8	0.9
Industrial	1.8	1.9–2.0
Miscellaneous	8.7	9.0–9.5

tests are performed concerning the mechanical behavior of salt under the stresses of unequal and changing loads. Crude oil and processed hydrocarbons are conveniently stored in the caverns produced in these domes by solution mining of the salt. The hydrocarbons float on the brine within the cavern and are readily recovered by simply pumping brine back into the well. The Gulf Coast, with its salt and hydrocarbon deposits, deep water ports, and ready sources of energy, is ideally suited for the chemical manufacture of salt derivatives, especially chlorinated hydrocarbons.[2]

Artificial brines also are produced extensively from stratified salt deposits in Michigan, Ohio, New York, West Virginia, Ontario, and Western Canada. Much of the early chemical industry in North America was concentrated in these areas because of the availability of these brines and the need for salt in the production of soda ash, chlorine, and caustic. The stratified salt exists in layers 20 to 200 feet thick, associated with anhydrite ($CaSO_4$). The nature of these deposits is such that it leads to higher calcium levels in the brine when recovered by solution mining compared with the brine obtained from salt domes. In the vicinity of many of these stratified salt deposits, there also exist naturally occurring brines, which consist mainly of $CaCl_2$, with lesser amounts of $MgCl_2$, KCl, and NaCl, a few thousand parts per million of bromide and up to a few hundred parts per million of iodide. These natural brines are mined principally for their bromine content.

Salt also is derived from seawater in those places in the world, such as California, where solar evaporation is sufficient to concentrate the seawater in large basins until some of the salt deposits. The salt-based chemical industry in Asia and the West Coast of the United States derives its salt from solar evaporation, which accounts for about half the total world production (see Table 12.2).

TABLE 12.2 U.S. Salt Production in 1987 (Millions of Short Tons) and Unit Value of Shipments[3]

Evaporated salt	5	($93.00/ton)
Rock salt	12	($14.75/ton)
Brine	20	($4.50/ton)
Total	37	

Salt and other chemicals also are derived from naturally occurring brines, such as brine from Searles Lake in California, which typically contains 16.5 percent NaCl, 6.8 percent Na_2SO_4, 4.8 percent KCl, 4.8 percent Na_2CO_3, and 1.5 percent $Na_2B_4O_7$. Another source is the Great Salt Lake, with 27 percent salinity, of which 80 percent is NaCl.

At Searles Lake, the brine is concentrated in a multi-effect evaporator, from which NaCl is recovered. The process liquor is fractionated further to yield a KCl–borax concentrate and burkeite, a double salt of Na_2CO_3 and Na_2SO_4, which is processed still further to yield soda ash and salt cake (Na_2SO_4) as products.

The use of salt for ice control on highways is second only to the salt usage for caustic chlorine manufacture. In order to minimize damage to the environment, Germany has practiced the application of damp salt on highways. In this process, solid NaCl is dampened with $CaCl_2$ solution, which prevents the salt from being blown from the road and initiates the thawing process more quickly.

SODA ASH

Soda ash, sodium carbonate, is derived commercially from three sources, which are, in order of decreasing importance: (1) naturally occurring trona ore; (2) the Solvay ammonia-soda process; and (3) naturally occurring

Fig. 12.1. U.S. production of synthetic vs. natural soda ash.[3]

alkaline brines. There has been a dramatic shift away from the synthetic or Solvay process ash in recent years, as is shown in Fig. 12.1.[3] No synthetic sodium carbonate has been produced in the United States since 1986. The Solvay process is still a major factor in ash production in Europe, but may decline there as well.

Like its predecessor, the Leblanc process, the Solvay process is on the decline for economic reasons. Increasing costs of production and environmental factors are the key issues in the many closings of synthetic ash plants in the last several years. However, a complete worldwide takeover by natural ash is doubtful because of two factors: (1) the limited amount of natural ash compared with the widespread availability of salt and limestone (the essential ingredients of the synthetic ash process); and (2) the locations of natural ash deposits relative to the locations of the ash consumers. Most of the U.S. natural ash is derived from the area of Green River, Wyoming. Significant amounts also are recovered from the alkaline brines of Searles Lake, California.

The synthetic ash process, or ammonia-soda process, was perfected by Ernest Solvay in 1865. The process is based on the precipitation of $NaHCO_3$ when an ammoniated solution of salt is carbonated with CO_2 from a coke-fired lime kiln. The $NaHCO_3$ is filtered, dried, and calcined to Na_2CO_3. The filtered ammonium chloride process liquor is made alkaline with slaked lime, and the ammonia is distilled out for recycle to the front end of the process. The resultant calcium chloride is a waste or by-product stream. The net overall chemical change in the process (shown schematically in Fig. 12.2) is represented by the following stoichiometry:

$$CaCO_3 \text{ (limestone)} + C \text{ (coke)} + O_2 \text{ (air)} + 2NaCl \text{ (brine)} \rightarrow Na_2CO_3 + CO_2 + CaCl_2 \text{ (brine)}$$

The process requires a large amount of fuel to calcine both the limestone and sodium bicarbonate and to generate steam for ammonia recovery. For the reaction proper, no fuel is required. In fact, large volumes of cooling water are required to remove the heat generated by the absorption and reaction of ammonia and carbon dioxide. The process has an imperfection in that an undesirable solution of calcium chloride also is produced.

Prior to being fed to the process, the NaCl brine must be purified to remove calcium and

Fig. 12.2. Simplified diagrammatic flow sheet for the Solvay ammonia-soda process.

magnesium ions so that they will not be precipitated when the brine is carbonated, producing objectionable scale on equipment surfaces, as well as contaminating the product. A solution of soda ash and caustic is added to the brine to precipitate the calcium as $CaCO_3$ and the magnesium as $Mg(OH)_2$. These impurities are flocculated, removed in a settler, and discarded. The purified brine is fed to the ammonia absorber in the process.

Lime for recovering ammonia in the process and CO_2 for reacting with the ammoniated brine are produced by calcining the best available limestone with foundry coke in vertical shaft kilns. This type of equipment is preferred for producing the maximum yield of active lime and the maximum concentration of CO_2 in the kiln gas. The dry lime, drawn from the vertical kiln, is cooled by the entering air, which is thereby preheated. Likewise, for maximum fuel economy, the exiting gas preheats the incoming limestone and fuel. The gas composition should exceed 40 percent CO_2 with only factional percentages of CO and O_2.

The dry lime and hot water are fed to a rotating cylindrical slaker to produce milk of lime, which is pumped to the distillers for the recovery of ammonia.

The Solvay process recycles large quantities of NH_3, and it is necessary to minimize losses in scrubbing the various NH_3-containing gas streams. Thus, the incoming purified brine is used to wash the ammonia-bearing air, which is pulled through the bicarbonate cake on the vacuum filters. This is accomplished in a packed absorber. The brine then passes through a second packed section where it absorbs the NH_3 in the gases from the carbonating towers. The brine then flows to the main NH_3-absorber, which is also a packed tower. The brine is circulated through water-cooled heat exchangers to remove the heat of absorption. Ammonia, carbon dioxide, and small amounts of H_2S released in the distiller are absorbed in the packed tower. In addition, a small amount of makeup ammonia is added. A typical analysis of the cooled, ammoniated brine is:

Temperature	38°C
NH_3	90 g/l
CO_2	40 g/l
NaCl	260 g/l
H_2S	0.1 g/l

The total heat removed from the ammonia absorption is 1.25 million Btu per net ton of soda ash produced.

Next, the ammoniated brine is carbonated to a point just short of crystallization. Then the brine is given a final carbonation and cooled to produce the crude $NaHCO_3$ crystals. The equipment used in the carbonation step consists of groups of five identical towers having alternate rings and discs in the upper section to assure mixing of the falling liquor with the rising gas stream without plugging by the crystallizing solid phase. The lower section of each tower consists of a series of heat-exchanger bundles alternating with rings and discs.

The ammoniated brine is passed downward through one of the group of five columns, which has become fouled with sodium bicarbonate after operating as a crystallizing unit for several days. The 40 percent CO_2 gas from the kilns is pumped into the bottom of the column to provide agitation and heat in order to dissolve the crystalline scale and bring the liquor to a composition just short of crystallization. The liquor is adjusted to the desired temperature by passage through a heat exchanger in preparation for feeding to the crystallizing towers, and then it is fed into the top of each of the other four columns in the group.

A mixed gas of 60 to 75 percent CO_2 derived from mixing the 40 percent CO_2 from the kilns and the 90 percent CO_2 from the calcination of bicarbonate, is fed to the bottom of these crystallizing units. Absorption of CO_2 in the highly alkaline ammoniated brine results in crystallization of crude sodium bicarbonate. Because of the heat evolved in the absorption and neutralization of the carbonic acid gas and from crystallization of the sodium bicarbonate, the temperature of the liquor in the column rises from 38°C to a maximum of about 62 to 64°C. In normal operation, the temperature of the discharge slurry is maintained at about 27°C by automatic adjustment of the water flow through the cooling tubes. A tower such as that shown in Fig. 12.3 has a capacity for producing 50 tons/day of finished soda.

Fig. 12.3. A carbonating tower in the ammonia-soda process, 69 feet high and 6 feet in diameter. (1) Entry for ammoniated brine, used when the tower is being cleaned; (2) entry for the ammoniated brine for the regular bicarbonate precipitation; (3) and (4) carbon dioxide entries; (5) outlet for bicarbonate slurries; (6) cooling water inlet; (7) cooling water outlet; (8) escape for uncondensed gases. (*Modeled after Kirschner.*)

The heat removed from the carbonator is about 260,000 Btu/net ton of ash produced, whereas that removed from the crystallizing units is 1.25 million Btu/ton of product.

It is noteworthy that the presence of sulfide in the feed liquor to the towers serves to maintain a protective film of iron sulfide on the cast iron equipment, which minimizes the contamination of product crystals from iron corrosion products.

Slurry drawn from the crystallizing columns is filtered in rotary vacuum filters where the $NaHCO_3$ crystals are water-washed. The filtered liquor then flows to the distillation unit for recovery of ammonia. The crude bicarbonate filter cake contains 3 percent NH_4HCO_3, 12 to 15 percent H_2O, and 80 percent $NaHCO_3$.

Next, the liquor is pumped to the NH_3-still preheater. At this point, the sulfide solution required for corrosion protection is added. The preheated process stream enters a stripper where excess CO_2 is removed prior to treatment with lime and distillation removal of the NH_3. The hot liquor from the stripper flows to an agitated vessel, to which is added milk of lime to release NH_3:

$$Ca(OH)_2 + 2NH_4Cl \rightarrow CaCl_2 + 2NH_3 + 2H_2O$$

The lime-treated solution then is fed to the top of the bubble-cap distillation unit. Steam is injected at the bottom, stripping out the ammonia down to a residual level of only 0.001 percent.

The crude $NaHCO_3$ is calcinated in dryers constructed with rotating seals and gas-tight feed and discharge mechanisms, to ensure the production of CO_2 that is undiluted with air:

$$2NaHCO_3 \rightarrow Na_2CO_3 + H_2O + CO_2$$

The heat requirement is about 2 million Btu per ton of soda ash produced. Product from the dryers is cooled for shipment, or converted to other products. A typical analysis of good commercial light soda ash made by this process is:

Na_2CO_3	99.70%
NaCl	0.12%
H_2O	0.12%
Fe	15 ppm
Ca and Mg	75 ppm

TABLE 12.3 U.S. Consumption of Sodium Carbonate (Thousand Short Tons)[3]

Total exports	2224
Glass	3475
Chemicals	1300
Pulp and paper	200
Water and gas stack treatment	450
Soaps and detergents	650
Miscellaneous	515
Total	8814

The waste liquor from the process may be evaporated to produce by-product $CaCl_2$ and NaCl. However, the major portion is clarified and pumped into water courses whose natural flow is sufficient to provide the dilution needed for disposal.

Production of soda ash from trona ore mined in the Green River area of Wyoming is rapidly expanding. Two basic processes are used, involving dry mining of the ore from a depth of about 1500 feet. A third process based on solution mining of the ore has been developed. In the older dry-mining process, the ore is crushed and dissolved at the surface. Then the solution is purified by settling and filtration, followed by evaporative crystallization to form sesquicarbonate: $Na_2CO_3 \cdot NaHCO_3 \cdot 2H_2O$. The crystals are centrifuged off and calcined to ash in steam-tube dryers. In a more recent process the crushed ore is calcined immediately, then dissolved, purified, and evaporatively crystallized to sodium carbonate monohydrate, which is centrifuged and dried to product ash.

Relatively small quantities of soda ash are produced from alkaline brines at Searles Lake, California, by a process of fractional crystallization that also produces other sodium and potassium salts. Table 12.3 shows the current distribution of soda ash in the United States.

The 1990 contract price of soda ash at plants in Wyoming was about \$93/ton.[6]

SODIUM BICARBONATE

Sodium bicarbonate is produced from a saturated solution of sodium carbonate by

TABLE 12.4 U.S. Consumption of Sodium Bicarbonate (1989) (Thousand Short Tons)[4]

Food	130–140
Animal feeds	85–90
Chemicals	40
Cleaning products	30
Pharmaceuticals/toiletries	20–25
Fire extinguishers	10
Other	35–40
Total	350–375

carbonation:

$$Na_2CO_3 + CO_2 + H_2O \rightarrow 2NaHCO_3$$

The precipitated bicarb crystals are centrifuged and then dried at low temperature to avoid reversion to carbonate.

Sodium bicarbonate also can be obtained by solution mining of natural sodium bicarbonate (nahcolite). In the mining procedure, pairs of well are sunk, warm water is pumped down one well, and dissolved sodium bicarbonate solution is pumped up from the other well. Solid sodium bicarbonate subsequently is crystallized from the solution.

The major use of sodium bicarbonate in North America is in foods as baking soda and as feed supplement for dairy cattle. Table 12.4 gives a breakdown of uses of $NaHCO_3$. The 1990 list price of $NaHCO_3$ was about 19 cents/lb.

SODIUM SULFATE

Production capacity for sodium sulfate, or salt cake, has remained fairly constant over the last decade, with the number of producing plants continuing to decline. Over half of the U.S. supply of Na_2SO_4 is produced from natural brines. Most of the remainder is by-product derived from the production of viscose rayon, sodium dichromate, and sulfonation process phenol. Salt cake produced in the manufacture of HCl from salt by the action of sulfuric acid (Mannheim process), or SO_2 and air (Hargreaves process), is on the decline.

TABLE 12.5 1986 Sodium Sulfate Production by Process[5]

Process	*Thousands of Tons*	*Percent*
Natural brine	396	48
Viscose rayon	136	17
Mannheim, Hargreaves	114	15
Dichromate, phenol, and other	172	21

Table 12.5 gives the 1986 U.S. production of salt cake. In addition, a small amount of Glauber's salt ($Na_2SO_4 \cdot 10H_2O$) still is produced for sale, but it has very little effect in the world of sodium sulfate.

The term "salt cake" now refers generally to sodium sulfate, but originally meant sulfate produced in the Mannheim furnace, which yields a product containing less than 99 percent Na_2SO_4, compared to the much purer anhydrous Na_2SO_4 derived from natural brines.

In the latter process (shown schematically in Fig. 12.4) the natural brine, containing about 10 percent Na_2SO_4, is saturated with NaCl by pumping it into a salt deposit. The concentrated brine then is pumped from the salt well through a refrigerated heat-exchanger, where it is cooled to around $-10°C$, and then into a crystallizer where quite pure crystals of Glauber's salt are formed. This solid hydrate of Na_2SO_4 then is melted and the water removed by evaporation using a submerged burner as the heat source. The wet Na_2SO_4 is dried further in a rotary kiln, producing anhydrous Na_2SO_4. Over 10 tons of natural brine and 1.6 million Btu of heat energy are required to produce one ton of product. The most significant production location is in Texas, where the sulfate-containing brine is conveniently located near deposits of both domed salt and natural gas.

A similar process is practiced at Searles Lake. There the brine first is carbonated and chilled to remove sodium carbonate and borax. Further chilling crystallizes Na_2SO_4 as the hydrate, Glauber's salt, and some remaining borax. The coarse crystals of Glauber's salt are separated from the fine

Fig. 12.4. Simplified flow sheet for the production of Na_2SO_4 from natural brine.

crystals of borax in a hydraulic classifier. The sulfate fraction then is filtered, washed, dried, and evaporated to produce anhydrous Na_2SO_4.

In another process at Searles Lake, the brine first is evaporated to produce NaCl and a double salt of Na_2CO_3 and Na_2SO_4. The two solids are separated in a hydraulic classifier, and the sulfate fraction is redissolved and recrystallized to Glauber's salt. Anhydrous Na_2SO_4 is recovered by mixing the Glauber's salt with sodium chloride brine. This dissolves the Glauber's salt and allows anhydrous Na_2SO_4 to precipitate because of its lowered solubility in the presence of sodium chloride.

In the Mannheim process, NaCl and 100 percent H_2SO_4 in amounts equivalent to complete conversion to Na_2SO_4 are fed to a circular muffle furnace made of cast iron. The furnace is equipped with a shaft, which penetrates from below and carries four arms, each of which is fitted with two cast iron plows. The shaft rotates at 0.5 rpm, slowly plowing the heated mixture to the circumference where the burned cake discharges through a chute. The furnace charge is heated to about 840°C, just below the fusion temperature of the salt cake. HCl is liberated, cooled, and absorbed in water to produce 32 percent HCl for sale. The salt cake is really a by-product of the operation.

A substantial amount of by-product Na_2SO_4 also is produced in a variety of processes involving the use of H_2SO_4 to react with a sodium salt, or to neutralize free caustic soda. Chief among these are the viscose rayon process and the manufacture of sodium dichromate. In the viscous process, Na_2SO_4 is formed by the reaction between H_2SO_4 and cellulose xanthate plus free caustic in the rayon fiber spinning bath. Spin-bath liquor is evaporated to crystallize Glauber's salt, which is centrifuged off, melted, and then evaporated to anhydrous Na_2SO_4. This is a significant source of sulfate, as 1.1 lb of Na_2SO_4 is produced per pound of rayon fiber spun.

In the manufacture of sodium dichromate, by-product Na_2SO_4 crystallizes directly as the anhydrous material when H_2SO_4 is added to a boiling solution of sodium chromate:

$$2Na_2CrO_4 + H_2SO_4 + H_2O \rightarrow Na_2Cr_2O_7 \cdot 2H_2O + Na_2SO_4$$

Decreased use of chrome by automobile and small-appliance manufacturers and mandated recycling of chromium wastes in the electroplating industry have led to a decline in Na_2SO_4 produced from this source.

Table 12.6[6] shows the U.S. distribution for consumption of Na_2SO_4. The price of salt cake is \$65/ton and for the higher-quality detergent and rayon grade \$90 to \$100/ton.

In North America, Canada is a net exporter of sodium sulfate, the majority of the exported

TABLE 12.6 U.S. Distribution of Na_2SO_4[6]

Industry	*Thousands of Tons*	*Percent*
Kraft pulp	275	30
Detergents	460	50
Glass	65	7
Miscellaneous	114	13

product going to the United States. Canadian sodium sulfate production is about 450 metric tons/year, and about 85 percent of this is produced from natural brine in Western Canada.

In both Canada and the United States the major end use of sodium sulfate is in kraft pulping. In the kraft pulping process, sodium sulfate (salt cake) is added as makeup to the recycle black liquor just prior to entering the recovery furnace. In the recovery furnace the sodium sulfate is reduced to sodium sulfide. A decade ago, sulfur losses accounted for a sodium sulfate makeup of about 80 lb/ton of pulp. Since then, antipollution regulations have encouraged technical innovations to decrease sulfur losses in the pulping process to a level where a makeup of only 40 lb salt cake/ton of pulp is required. Also, in recent years the use of Na_2SO_4 in the manufacture of kraft paper has been giving way somewhat to the use of sodium sulfide, sodium hydrosulfide, and emulsified sulfur, which are more readily available as by-products of the petroleum industry, derived from the caustic scrubbing of sulfide-containing hydrocarbon gases. The sulfides often are preferred by the paper industry because the sulfate must be reduced to sulfide in the pulping process.

In dry detergents, sodium sulfate is used largely as a low-cost inert filler and diluent. In recent years detergents have become more concentrated, and sodium sulfate requirements have been reduced.

In glass manufacture, sodium sulfate is used in small amounts for its sulfur and sodium oxide values as a melting aid. In order to minimize the deleterious effects of volatilized sodium sulfate from glass manufacture, the amount of sodium sulfate consumed per ton of glass has been decreasing in recent years.

SODIUM SULFIDES

Sodium sulfide (Na_2S) and sodium hydrosulfide, sometimes referred to as sodium sulfhydrate (NaHS), share the same derivation: caustic soda and H_2S. Earlier processes based on reduction of salt cake with coal have largely been supplanted. In 1990 the U.S. installed capacity for NaHS was 134 tons/year. In addition, more than 100,000 tons/year of lower-purity lower-strength sodium hydrosulfide was recovered from small refineries, bromine operations, and other sources. The source of H_2S used in the manufacture of NaHS is essentially all by-product. Controlled caustic scrubbing of H_2S forms a solution of NaHS:

$$NaOH + H_2S \rightarrow NaHS + H_2O$$

The NaHS solution is filtered to remove the sulfides of heavy metals, such as Fe, Hg, Ni, Mn, and Cu. The clear filtrate may be sold as a 44 to 46 percent solution of NaHS or evaporated in stainless steel equipment to crystallize a solid hydrate containing 70 to 72 percent NaHS, which is sold as a flake product.

NaHS is easily converted to Na_2S by further reaction with caustic:

$$NaHS + NaOH \rightarrow Na_2S + H_2O$$

By using a NaHS solution of the proper concentration with a flake caustic, a hydrate product containing 60 to 62 percent Na_2S is obtained. This is sold directly as a flake product, or fused solid in drums. When high-quality H_2S and NaOH are used, the product Na_2S is suitable for use in dyes, photography, rayon, and leather manufacture.

Lower-quality sulfides are obtained by using H_2S produced as a by-product in the manufacture of CS_2 from methane (or other low molecular weight hydrocarbons) and sulfur:

$$CH_4 + 4S \rightarrow CS_2 + 2H_2S$$

The gas mixture from this catalytic reaction is cooled and scrubbed with caustic. The sulfides so produced contain small amounts of mercaptans (e.g., CH_3SH), which lend a very objectionable odor to the product.

TABLE 12.7 Uses for Na_2S and NaHS[6]

Industry	*% Total Market* Na_2S	*NaHS*
Pulp and paper	3	36
Dyestuffs and misc. chemicals	29	40
Rayon and film	3	6
Metals and minerals	3	—
Leather	62	18

Sulfides also are obtained as by-products in the manufacture of $BaCO_3$ from the barite ore, $BaSO_4$, by roasting with coal, water leaching, and treatment with soda ash:

$$BaSO_4 + 4C \rightarrow BaS + 4CO$$

$$BaS + Na_2CO_3 \rightarrow Na_2S + BaCO_3$$

Table 12.7 shows the distribution for consumption of Na_2S and NaHS by industry use. They both are used extensively in a variety of industries, and the use of NaHS in paper pulping is growing. The 1990 carload price for NaHS for 45 percent liquid (100% basis) was \$500/ton. Tight caustic soda supplies have led to higher costs for an industry already faced with weak prices and overcapacity.

SODIUM THIOSULFATE

Eighty percent of sodium thiosulfate (hypo) ($Na_2S_2O_3$) is used in photography (Table 12.8) because of its ability to dissolve water-insoluble silver salts. However, because substantial market erosion has occurred concurrently with the growth in use of videotape, the market for sodium thiosulfate continues a steady decline from 40,000 tons/year 30 years ago to well under half that today.

TABLE 12.8 Sodium Thiosulfate Major Uses[6]

Industry	*Percent*
Photography	80
Oil recovery	7
Leather	5
Water treatment	5
Other	3

Most current production is derived as by-product from the manufacture of sulfur dyes and Na_2S:

$$Na_2S + SO_2 + H_2O \rightarrow Na_2SO_3 + H_2S$$

$$Na_2SO_3 + S \rightarrow Na_2S_2O_3$$

$$2Na_2S + Na_2CO_3 + 4SO_2 \rightarrow 3Na_2S_2O_3 + CO_2$$

In producing sulfur dyes, by-product leach liquor contains sodium thiosulfate, which is simply derived by evaporation and crystallization. Much of the thiosulfate is sold as the pentahydrate, $Na_2S_2O_3 \cdot 5H_2O$. Stainless steel equipment is used in processing thiosulfate.

Of less importance today is the older soda ash–sulfur process. Ash is dissolved in hot water, and the solution is pumped to SO_2 absorption towers. The resulting $NaHSO_3$ liquor is then heated with powdered sulfur in an agitated stainless steel digestion tank at elevated temperature. The product solution is filtered, evaporated, and crystallized to yield $Na_2S_2O_3 \cdot 5H_2O$ product, which is centrifuged, washed, dried, and packaged in airtight containers (to discourage efflorescence). In 1990 sodium thiosulphate pentahydrate sold for 32 cents/lb.[6]

SODIUM SULFITE

Sodium sulfite (Na_2SO_3) is produced by reacting sulfur dioxide with a solution of sodium carbonate or sodium hydroxide. Sodium sulfite also is recovered as a by-product of resorcinol production. About half of the sodium sulfite produced in North America is used by the pulp and paper industry (i.e., chemithermomechanical process). It also is used in photography, textile bleaching, and food preservation, and as a chemical intermediate. In 1990, 160,000 short tons of sodium sulfite were produced in the United States, excluding material produced and used captively by some paper mills. The 1990 selling price for sodium sulfite (anhydrous 95–100%) was about 24 cents/lb.

SODIUM BISULFITE

Most of the sodium bisulfite of commerce is really the anhydride, $Na_2S_2O_5$, sodium metabisulfite or sodium pyrosulfite. Sodium bisulfite is not stable and is generally transported in the metabisulfite form, which is a solid containing 98 percent $Na_2S_2O_5$, 1.5 percent Na_2SO_3, and 0.5 percent Na_2SO_4. Methods of manufacture of sodium metabisulfite are all variations of the same theme:

$$Na_2CO_3 + 2SO_2 \rightarrow Na_2S_2O_5 + CO_2$$

In one variation, SO_2 is sparged into a stainless steel absorber through which is passed a solution of Na_2CO_3. Formation of product crystals from the saturated solution is achieved by lowering the temperature. The crystals are centrifuged off and dried rapidly in a flash dryer in order to avoid air oxidation to sulfate. The markets for sodium metabisulfite include chemical intermediates, pharmaceuticals, food preservatives, and dye fixing. Additionally, sodium metabisulfite is used in photography, as an antichlor to remove excess chlorine from bleached kraft pulps, and for the manufacture of sodium hydrosulfite. North American sodium metabisulfite production in 1988 for merchant sales was about 50,000 tons, and the 1990 selling price for the anhydrous product was 28 to 32 cents/lb.

SODIUM HYPOSULFITE (HYDROSULFITE)

Not to be confused with "hypo" (the term used in photography to refer to thiosulfate), sodium hyposulfite (or hydrosulfite) ($Na_2S_2O_4$) also is referred to as sodium dithionite. $Na_2S_2O_4$ is a powerful reducing agent used principally for the reduction of vat dyes and brightening of mechanical wood pulps and clays (Table 12.9). There are numerous methods for production of sodium hyposulfite, all of which involve reduction of SO_2 or sodium bisulfite. Sulfur dioxide or $NaHSO_3$ is reduced by Zn, sodium amalgam, sodium formate, sodium borohydride, or cathodic reduction to produce $Na_2S_2O_4$.

TABLE 12.9 1990 U.S. Distribution of Merchant Sodium Hydrosulfite[4]

	Percent
Pulp bleaching	33
Textile dying	22
Clay bleaching	13
Exports	28
Other	4

There are two variations of the zinc method. In the first, a solution of sodium hydrogen sulfite is reduced with zinc dust in the presence of excess SO_2:

$$2NaHSO_3 + Zn + SO_2 \rightarrow Na_2S_2O_4 + ZnSO_3 + H_2O$$

Milk of lime is added to neutralize remaining sulfurous acid, and $CaSO_3$ and $ZnSO_3$ are removed by filtration. Common salt is added to the remaining solution to salt out $Na_2S_2O_4 \cdot 2H_2O$. The suspension of crystals is heated to 60°C to dehydrate the product, water is extracted with alcohol, and the product crystals are vacuum-dried. The crystals are stable only when completely dry.

The second method based on zinc consists of treating an aqueous suspension of zinc dust with SO_2 at 80°C:

$$Zn + 2SO_2 \rightarrow ZnS_2O_4$$

Treatment of this solution with soda ash precipitates a basic zinc carbonate, which is filtered off. The dry crystalline product $Na_2S_2O_4$ is obtained by salting out and drying, as in the first method. The zinc methods have become obsolete because of the environmental concerns associated with zinc discharge.

The amalgam process uses a dilute sodium amalgam (0.5% Na) derived from the cathode reaction of a mercury chlorine cell.[8] The amalgam is admixed with a solution of $NaHSO_3$ whose pH is controlled in the range of 5 to 7:

$$4NaHSO_3 + 2NaHg \rightarrow Na_2S_2O_4 + 2Na_2SO_3 + 2H_2O + Hg$$

The mercury is returned to the mercury cell to form more amalgam, and the $Na_2S_2O_3$ is allowed to react with aqueous SO_2 to form more $NaHSO_3$.

One common method for sodium hyposulfite production involves the use of sodium formate:

$$HCOONa + NaOH + 2SO_2 \rightarrow Na_2S_2O_4 + H_2O + CO_2$$

The above aqueous reaction requires the use of methanol under acidic conditions. In this process, the anhydrous sodium hydrosulfite is directly precipitated and is filtered off.

Sodium hyposulfite (5–7% $Na_2S_2O_4$) also is produced on-site at pulp and paper mills by reduction of SO_2 with sodium borohydride solution under alkaline conditions (Borol process):

$$\underbrace{(NaBH_4 + 3.4NaOH)}_{\text{Borol solution}} + 5.6NaOH + 9SO_2 \rightarrow 4Na_2S_2O_4 + NaBO_2 + NaHSO_3 + 6H_2O$$

In a recent patent[9] an electrochemical process for producing hydrosulfite solutions was claimed. The cathode reaction is:

$$2HSO_3^- + 2H^+ + 2e^- \rightarrow S_2O_4^{2--} + 2H_2O$$

It is anticipated that such a process would be viable for on-site production of sodium hyposulfite.

In 1990 there were about 80,000 tons of merchant sodium hyposulfite produced. In addition, about 10,000 tons of sodium hyposulfite were produced on-site at pulp mills by reacting sodium borohydride and SO_2. The on-site process has the advantage of minimum hyposulfite degradation to thiosulfate, which is known to cause corrosion problems in pulp mills. The 1990 list produce for $Na_2S_2O_4$ was about 73 cents/lb, but discounting from the list price was prevalent.

SODIUM PHOSPHATES

In 1990 little over half of the 320,000 tons of phosphorus produced in the United States was utilized in the manufacture of the family of sodium phosphate products, chief among which is sodium tripolyphosphate, $Na_5P_3O_{10}$. Sodium tripolyphosphate (STPP) is approximately 57 percent P_2O_5 and 42 percent Na_2O. In 1990, 500 tons of STPP was produced and sold at about 30 cents/lb. It is made by the reaction between soda ash and phosphoric acid.[10] By far, the major source of phosphoric acid is furnace-grade acid, because of its high purity as compared with the less pure wet-process acid. A hot solution of Na_2CO_3 is reacted with 60 percent H_3PO_4, with a molar ratio of sodium to phosphorus of 1.67 maintained in order to achieve the desired stoichiometry. The solution is purged with steam to evolve all the CO_2, and then it is filtered, spray-dried, and granulated for shipment. By changing the ash-to-acid ratio, a variety of products can be produced.

This process forms the basis for an entire family of products, fundamental to which, in addition to the tripolyphosphate, is the production of the di- and trisodium salts. They may be produced in the same process by crystallizing out and recovering $Na_2HPO_4 \cdot 12H_2O$.

The major use of STPP is in home laundry, industrial, and institutional detergents, as indicated in Table 12.10.[6]

Phosphate bans in detergents continue to be legislated. The total demand for STPP in the last few years has declined at a rate of about 1.5 percent per year. The primary reason for this decline has been the penetration of the phosphate-free liquid detergents into the heavy-duty phosphate detergent powder market.

The mono-, di-, and pyrophosphates are used in foods, the latter two in baking powders. The monosodium salt is produced

TABLE 12.10 1990 STPP Uses[6]

Industry	*Percent*
Home laundry detergents	44
Industrial/institutional detergents	19
Dishwashing detergents	18
Food	4
Miscellaneous	2
Exports	13

from Na_2HPO_4 by further processing with H_3PO_4, evaporation, and crystallization. Prolonged heating of the monosodium derivative at 230°C produces the pyrophosphate, $Na_2H_2P_2O_7$. Further heating of the monosalt to the point of fusion produces hexametaphosphate, a glassy substance of uncertain stoichiometry, which is used in water treatment.

SODIUM SILICATE

The sodium silicate family is analogous to the phosphates in that a whole series of derivatives are produced in a reaction between Na_2CO_3 and silica by varying the ratio of soda ash to silica. The products often are referred to by the ratio of SiO_2 to Na_2O in a given composition. The U.S. production was almost 900,000 tons in 1990. Table 12.11 gives a typical use pattern for this heavy chemical group.

Sodium silicates are made in batches by fusing sand and soda ash in a furnace at about 1300°C, in whatever ratio is necessary to achieve the desired properties of the final compound. In general:

$$Na_2CO_3 + nSiO_2 \rightarrow Na_2O(SiO_2)_n + CO_2$$

$$2NaOH + nSiO_2 \rightarrow Na_2O(SiO_2)_n + H_2O$$

The product is called water glass, because, when solid, it actually is a glass; but, unlike lime-soda glass (ordinary window glass), it is soluble in water. The process is carried out in large tank furnaces similar to window glass furnaces. The materials are introduced in batches at intervals, but the products may be drawn off continuously if desired. A mixture of salt cake and coal may replace a portion of the soda ash.

As the melt leaves the furnace, a stream of cold water shatters it into fragments. These are dissolved with superheated steam in tall, narrow steel cylinders with false bottoms,[11] and the product liquor is clarified.[12] Sodium silicates are sold as solutions that vary from the most viscous, 69°Be, to the thinner solutions, down to 22°Be, suitable for use in paints. The composition of typical commercial grades of sodium silicate solutions is shown in Table 12.12. A dry powder also is marketed (Table 12.13); it is made by forcing the thick liquor through a very fine opening into a chamber swept with a rapid current of cold air, which carries off the moisture.[13] Because the solid silicates are hygroscopic, they often are blended with Na_2SO_4 to prevent caking. Silicate solutions also are processed further with mineral acids to manufacture hydrated

TABLE 12.11 1990 Uses for Sodium Silicates[6]

Industry	*Thousands of Tons*	*Percent*
Soaps and detergents	261	29
Catalysts and gels	252	28
Pigments	198	22
Boxboard adhesives	72	8
Paper and ore treatment	48	6
Other		7

TABLE 12.12 Composition of Typical Commercial Sodium Silicate Solutions[6]

Weight Ratio SiO_2:Na_2O	*% SiO_2*	*% Na_2O*	*Solids Content (%)*	*Specific Gravity (°Be)*	*Viscosity at 20°C (Centipoises)*	*Density (Pounds per Gal)*
3.22	28.7	8.9	38.0	41.2	290	11.7
3.00	30.9	10.3	41.2	44.6	412	12.0
2.84	31.9	11.2	43.1	47.0	690	12.3
2.58	32.1	12.5	34.6	NA	780	12.6
2.54	26.9	10.6	37.5	42.0	63	11.7
2.40	33.4	13.9	47.0	52.2	1700	13.0
2.00	29.4	14.7	44.1	50.0	335	12.7
1.80	24.1	13.4	37.5	54.6	60	13.4

TABLE 12.13 Commercially Available Alkaline Sodium Silicate Solids[6]

Name	*Chemical Formula*	*Molecular Ratio* $Na_2O:SiO_2$
Sodium metasilicate, anhydrous	$Na_2O \cdot SiO_2$	1.0
Sodium metasilicate, pentahydrate	$Na_2O \cdot SiO_2 \cdot 5H_2O$	1.0
Sodium sesquisilicate	$3Na_2O \cdot 2SiO_2$	1.5
Sodium orthosilicate, anhydrous	$2Na_2O \cdot SiO_2$	2.0
Sodium orthosilicate, hydrate	$2Na_2O \cdot SiO_2 \cdot 5H_2O$	2.0

silica, which yields silica gel upon being dried. Silica gel is used extensively in air-drying equipment and in the manufacture of catalysts.

There continues to be an overcapacity for sodium silicate manufacture in Canada and the United States. The sharply growing use of peroxide bleaching and waste paper deinking is boosting sodium silicate demand in the pulp and paper industry, particularly in Canada. The silicate sequesters iron and other metal ions that reduce the peroxide efficiency. It also keeps the pH buffered in the alkaline range required for peroxide bleaching of pulp. Sodium silicate also has become an important ingredient of the chemical formulation for the deinking of printed waste paper. In this process, in addition to being a buffering and sequestering agent, silicate is known to provide detergency and anti-redeposition properties during the flotation deinking process.

In 1990[6] 41°Be solution with a 3.22 to 3.25SiO_2 : Na_2O weight ratio sold for about 10 cents/lb. The price for the similar-ratio solid product was about 22 cents/lb.

CHLOR-ALKALI (CHLORINE AND CAUSTIC SODA)

Ranking among the higher-volume chemicals produced in the world are chlorine and caustic (sodium hydroxide). In 1990 North America produced over 25 million tons of caustic and chlorine, and the value of these products was over $5 billion. Table 12.14 gives chlor-alkali production[4] in major regions of the world.

Historically, the electrolysis of brine accounts for approximately 96 percent of all chlorine production. About 2 percent of the total electrical power generated in the United States is used for chlor-alkali manufacture. Over a third of the total world chlor-alkali capacity exists in the United States in about 50 plants, by far the larger of which are concentrated in the Gulf Coast area of Louisiana and Texas.[14,15] This concentration derives from the combined availability in that region of salt, hydrocarbons (fuel and organic products based on chlorine and caustic), and deep-water ports for transportation. The intercoastal canal, the Mississippi and Ohio rivers, and their connecting waterways are also well suited for heavy barge traffic from the Gulf to much of the heartland.

TABLE 12.14 Chlor-Alkali Production in Major Regions of the World[4]

	Thousands of Short Tons	
	Caustic	*Chlorine*
U.S.A.*	12,572	11,915
Canada*	1,334	1,326
Western Europe	11,623	10,810
Eastern Europe	6,580	5,442
Japan	3,537	3,193

*1989 Statistics; others 1987.

Table 12.15 gives the distribution of U.S. chlorine by industry use. Table 12.16 similarly lists the market distribution for caustic.[6,16]

The heart of the chlor-alkali process is the electrolytic cell in which saturated, purified NaCl is electrolytically decomposed. There are basically three types of chlor-alkali cells currently in use, which are shown schematically

TABLE 12.15 Market Distribution of U.S. Chlorine[6,16]

Market	*Percent*
Ethylene chloride/vinyl chloride monomer	28
Pulp and paper	18
Inorganic chemicals	8
Chloromethanes	7
Chloroethanes	6
Propylene oxide and other organics	20
Water treatment	6
Miscellaneous	7

TABLE 12.16 Market Distribution of U.S. Caustic[6,16]

Market	*Percent*
Pulp and paper	22
Organic chemicals	20
Inorganic chemicals	11
Soaps and detergents	7
Petroleum	7
Water treatment	7
Textiles	5
Alumina	4
Other	9
Exports	8

in Fig. 12.5.[17–19] These cells are designed for continuous electrolysis of brine to produce chlorine gas, hydrogen gas, and aqueous caustic soda.

The diaphragm cell derives its name from the separator, or diaphragm, which is a mat of asbestos fiber vacuum-deposited directly on a woven-wire, mild steel member. The steel mesh serves doubly as a support for the diaphragm and as the cathode, or negative electrode, of the cell. The diaphragm separates the anode and cathode compartments of the cell to discourage direct mixing of the chlorine and caustic, and also to prevent chlorine and hydrogen from mixing, which can be a dangerous situation because they are capable of reacting violently with one another under certain conditions. A head of brine is maintained higher on the anode (the positive electrode, at which chlorine is produced) side of the diaphragm than the cell liquor level on the cathodic side. Thus brine slowly percolates through the diaphragm and into the cathode chamber of the cell, where water is decomposed to form hydrogen gas and hydroxide ion. The effluent cell liquor contains about 12 percent NaOH, 14 percent NaCl, and small amounts of chlorate, hypochlorite, sulfate, and carbonate—with the remainder being water. The usual caustic produced for sale is a 50 percent NaOH solution; so this cell effluent requires evaporation and separation of the salt.

The mercury cell differs from the diaphragm cell in that the cathode is a thin film of mercury, with no separator between it and the anode. Although chlorine is produced at the anode in both types of cell, the cathode reactions differ markedly. Instead of water decomposing at the cathode of the mercury cell, as it does in the diaphragm cell, sodium metal is deposited in the mercury to form a dilute (0.5%) amalgam. This amalgam flows out of the cell into a decomposer, in which water and the amalgam react directly to form a 50 percent NaOH solution of high purity, there being no evaporation required. The denuded mercury is returned to the cell.

Another difference between the two cell types is the brine flow. In the diaphragm cell about 50 percent of the NaCl is decomposed, and the remainder flows out of the cell in the weak caustic effluent. In the mercury cell, only a small fraction of the salt is decomposed per pass through the cell. The brine is immediately dechlorinated, resaturated with salt, and returned to the cell.

At first glance, it would appear that the mercury cell process is far superior to the diaphragm cell process. In many ways it is. However, the mercury cell requires much more electrical energy than does the diaphragm cell. In certain areas, especially where electrical energy is cheap and where small-capacity, easily run plants are required, the mercury cell is quite economical. Also, the mercury cell may be preferred if fuel costs are high, as that discourages the plant designer from installing evaporators—which require burning fuel just to make steam to run them. In large installations where it makes more sense to install a powerhouse at the site, the

Fig. 12.5. Chlor-alkali electrolysis cells: (a) diaphragm cell; (b) membrane cell; (c) mercury cell. (*Courtesy the Dow Chemical Company.*)[17]

generation of power cogenerates lower-pressure steam, which may be used quite economically to drive a multiple-effect evaporator.

The membrane cell has some operating features reminiscent of both the diaphragm and the mercury cell. This new cell is somewhat like a diaphragm cell, in that it has a separator. However, the similarity ends there. The separator in the newer type cell is a perfluorinated ion-exchange membrane, through which essentially only hydrated sodium ions migrate under the influence of the electrical field.[20,21] Because there is no fluid flow through the membrane, the brine in the anode compartment becomes depleted in salt. As it does so, it is removed, dechlorinated, resaturated, and returned to the cell, just as in the case of the mercury cell. Also, very high-quality caustic is made in this cell, in concentrations ranging from 30 to 35 wt. % NaOH with less than 50 ppm NaCl. These solutions are more economically evaporated than is the diaphragm cell effluent, because

Fig. 12.6. Chlor-alkali membrane cross section.[17]

Fig. 12.7. Perfluorinated polymers for chlor-alkali membranes.

they are more concentrated and contain no salt, sodium sulfate, or sodium chlorate impurity to speak of.[22] This technology has emerged in the last 10 years and in time will largely displace other cells.

As mentioned earlier, the ion-exchange membrane allows the transport of primarily sodium ions and water from the anolyte to the catholyte compartments, whereas the diaphragm is designed to allow percolation of all the anolyte through the separator. The cation-conducting, ion-exchange membrane is designed to reject anions, as indicated in Fig. 12.6. The chlor-alkali membranes in use today consist of one or more perfluorinated, ion-exchange polymeric materials, as illustrated in Fig. 12.7. The short-chain polymers shown in Fig. 12.7 have been developed by the Dow, Asahi Glass, and Asahi Chemical companies.[17] Long-chain polymers are manufactured by Du Pont and sold under the trade name Nafion.*

A chlor-alkali membrane, designed to produce 30 to 35 percent NaOH, consists of at least two distinctly different polymer layers, as shown in Fig. 12.6. The anode side of the membrane is about a 0.1 mm film of sulfonic acid polymer, whereas the cathode side is about a 0.05 mm film of carboxylic acid polymer. A Teflon* fabric net is embedded in the polymer film to provide mechanical strength to the membrane. In some cases, the membrane surface is modified to facilitate gas bubble release.

A most significant advance in cell anode technology is being used to great advantage in all three cell types to conserve energy, reduce impurities, and lengthen cell life.[23,24] Prior to this development the cell anodes were made out of graphite. The new anodes are made out of titanium, which has excellent corrosion resistance in hot, chlorinated brine. However, the very oxide film that lends corrosion resistance is not an electrical conductor. By applying electroactive coatings, largely composed of mixed oxides of ruthenium and titanium, to the surface of the titanium, it remains conductive. These anodes give a vast improvement in performance over graphite for two reasons: (1) they do not wear away like graphite, which, as it becomes increasingly thinner, widens the electrode gap in the cell, requiring higher and higher voltages to operate the cell; (2) the electroactive coating is actually a catalyst that allows chlorine to be formed on the anode surface with essentially zero activation energy. Both properties contribute to much reduced power consumption in running the cells. In the case of the mercury cell, there is the added advantage that the dimensional stability of

*Trademark of E. I. du Pont de Nemours & Company.

the new anode precludes the continual adjustment of electrode spacing, which is constantly necessary with graphite electrodes.

The ion-exchange membrane and dimensionally stable anode developments have impacted greatly on chlor-alkali operations, though the former is a very recent development. Their impact has been accelerated by both environmental and energy cost considerations. The cell capital cost is skyrocketing because of the expense of these new components; however, their usage does lower operating costs. In practice, a balance between capital costs and operating costs is sought to achieve an economic optimum for a given plant site.

Both membrane and diaphragm cells use punched plate, expanded mesh, or woven wire cathodes of nickel, nickel-plated steel, or stainless steel. In most modern membrane cells some type of activated or catalytic cathode is used. Usually the gap between the anode and cathode is minimized to reduce the electrical resistance of the cell.[17] Various cell configurations are shown in Fig. 12.8. A cell designed for zero gap operation usually has a compressible mesh to conduct the current. The SPE (Solid Polymer Electrode) concept shown in Fig. 12.8 is not commercial yet, but the concept is borrowed from fuel cell technology where the porous electrode is embedded into the membrane.

Fig. 12.8. Chlor-alkali cell configurations: (a) gap cell; (b) zero gap cell; (c) SPE cell.

The chemistry at the anode is identical for all three cells:

$$2Cl^- \rightarrow 2e^- + Cl_2$$

The cathode reactions also are the same for the diaphragm and membrane cells:

$$2H_2O + 2e^- \rightarrow H_2 + 2OH^-$$

In the mercury cell, however, the cathode reaction is quite different, there being no hydrogen formed in the cell:

$$Na^+ + e^- \xrightarrow{(Hg)} Na \text{ (amalgam)}$$

The amalgam may be used as a chemical reductant (i.e., to reduce $NaHSO_3$ to the hydrosulfite[8] or to hydrodimerize acrylonitrile to adiponitrile[25]), or decomposed with water to make caustic, as is generally done.

Environmental pressures on mercury discharges have dealt a blow to the practice of mercury cell technology. As a result, there have been many plant closings. An excellent job has been done in achieving sustantial reductions of Hg losses, and this technology will continue to be practiced. However, most recent emphasis is on the diaphragm cell and ion-exchange membrane technology. As the latter matures, further inroads will be made against the mercury cell. This has already happened in Japan.

The theoretical decomposition potential, $E°$, of a cell (i.e., the voltage below which it is impossible to form any products at all) is strictly dependent on the nature of the electrode reactions. It is related to the Gibbs free energy of the net overall chemical change occurring in the cell: $G = nFE°$, where n is the number of electrons involved in the electrode reaction, and F is the Faraday constant.[26] $E°$ is the thermodynamic potential, to which the IR drops (voltage necessary to drive the current through the electrical

resistances in the cell), and the overpotential (voltage associated with the kinetics of the reactions) must be added to get the actual cell voltage:

$$\text{Cell voltage} = \text{thermodynamic potential} + IR \text{ drop} + \text{Overvoltage}$$

Typical voltage components for a membrane cell are given in Table 12.17.

Using a modern diaphragm cell running at 1500 amps/m^2 as an example, the percentages of the total cell voltage attributable to these three components are: 65 percent thermodynamic, 25 percent *IR*, and 10 percent overvoltage. Cell design and development programs continue to search for better ways to reduce energy consumption. Rapidly rising power costs have made power consumption the primary criterion in evaluating cell performance.[27] Thus, higher cell capital becomes justifiable on the basis of performance. But as cell capital increases, it tends to force operation at increased current density in order to produce more pounds per cell per unit of time. In a practical sense, this translates to a compromise between cell cost and current density in order to minimize production cost. Figure 12.9 is a typical current–voltage curve for a diaphragm cell. The current density simply refers to total current divided by the nominal electrode area of either the anode or the cathode, often expressed as thousands of amperes (kA) per square meter. Figure 12.9 also shows the electrical energy consumption in kilowatt hours per 2000 lb of chlorine (kWh/T Cl_2).

Fig. 12.9. Current–voltage curve for a typical diaphragm chlor-alkali plant.

The minimum energy requirement, unattainable in practice, is 1543 kWh/T Cl_2. This corresponds to 100 percent current efficiency (685.8 kAh/T Cl_2) and 100 percent voltage efficiency (i.e., no *IR* drops and zero overvoltage at both electrodes, so that the voltage is the thermodynamic potential, or 2.17 volts). The energy efficiency of a typical diaphragm cell is easily expressed as:

$$\frac{(2.17) \times (\%\ \text{Current efficiency})}{(\text{Cell voltage})}$$

For example, if a cell running at 80,000 amperes at a current density of 2.2 kA/m^2 and a current efficiency of 96 percent requires a cell voltage of 3.40, the overall energy efficiency is:

$$\frac{2.17 \times 96}{3.40} = 61.3\%$$

Figure 12.10 shows the dramatic improvements made in the electrical energy consumption of membrane cells since the mid-1970s.[17]

Much of the inefficiency results in heat generation; a cell will heat itself almost to boiling when fed brine at about 40°C. The self-heating tends to lower the cell voltage because the electrical resistance of the cell decreases with increasing temperature. As modern rectification equipment (transformers, controllers, and rectifiers) is very efficient, the

TABLE 12.17 Typical Voltage Components for a Membrane Cell at 3.1 kA/m^2, 85°C, 34 Weight % NaOH, DSA Anode, 200 g/l NaCl Anolyte and Activated Cathode[16]

Component	*Voltage*	*% of Total*
Reversible potential of cell reactions	2.20	72
Membrane *IR* + concentration polarization	0.50	16
Electrode overpotentials	0.15	5
Solution *IR*	0.15	5
Structural *IR*	0.06	2
	3.06	100

Fig. 12.10. Electrical energy consumption improvements in membrane cells.[17]

conversion of AC electricity into electrochemical products is essentially dependent on the cell efficiency. Typically, AC-kWh/lb product = 1.03 × (DC-kWh/lb product).

Figure 12.11 is a schematic of a commercial diaphragm cell, typically operated in the range of 75,000 to 150,000 amperes. Many such cells are connected in series, as shown in Figure 12.12, so that the current leaving at the anode of one cell enters the cathode of the next cell, with the anode terminal cell connected to the positive bus of the rectifier, and the cathode terminal cell connected to the negative bus of the rectifier. Bipolar and monopolar cell circuit layouts are shown in Fig. 12.13.[17]

With allowance for inefficiencies, each 30 kA increment of current that passes through the series circuit of cells produces 2000 lb of chlorine, 2100 lb of caustic, and 57 lb of hydrogen, per cell, in 24 hours of operation.

If a production plant is required to produce 900 tons of chlorine per day using rectification equipment capable of sustained operation at 90,000 amperes, the plant must operate a minimum of 300 cells. If the cell voltage is 3.4 volts at 90 kA, rectification equivalent to 1020 volts is required. In practice, the logistics of production requirements, plant reliability, and maintenance outages will dictate the size (i.e., number of cells and nominal amperage) of each circuit. In the example above, the plant would use a minimum of two circuits, with two nominal 90-kA, 600-volt rectifiers.

Figure 12.14 is a simplified flow sheet for a typical diaphragm cell chlor-alkali plant. The process begins with brine treating, in which

Fig. 12.11. Diamond Shamrock diaphragm cell. (*Courtesy Diamond Shamrock Corporation.*)

the well brine is saturated with return salt from the caustic evaporators, then treated with hypochlorite to destroy ammonia and amines, and also treated with carb-caustic liquor to remove calcium and magnesium impurities. Membrane cells require a secondary purification of the brine feed to provide parts per billion levels of Ca, Mg, and other heavy metals.

The removal of trace ammonia-nitrogen impurities is essential to avoid buildup of NCl_3 in the chlorine purification and lique-

Fig. 12.12. A series of Diamond Shamrock cells. (*Courtesy Diamond Shamrock Corporation.*)

faction stages. The NCl_3 is unstable in concentrations above a few thousand parts per million, and can decompose violently. The Ca and Mg impurities are removed to reduce plugging of the cell diaphragms, and to lengthen the time between diaphragm washing or diaphragm replacement. This was typically done every few months, but has now been extended to over a year, principally by replacement of the graphite anodes with the newer metal anodes. In the older cells, the slowly oxidizing graphite formed a sludge that tended to plug the diaphragm.

The hydrogen from the cells formerly was discarded (vented to the atmosphere) unless it was used in some of the larger plants to make NH_3 and other chemicals. It takes a large plant to produce a useful amount of H_2 because the cell produces only 0.028 lb H_2 per pound Cl_2. Also, the fuel value of H_2 is only a third that of methane. But since the energy crisis of the 1970s, the fuel value of the hydrogen has become sufficient to recover it for boiler fuel. This requires cooling to remove water, scrubbing to remove traces of chlorine, caustic, and salt, and then compressing sufficiently to satisfy the feed requirements of the burner. Of increasing importance is the burning of the hydrogen with the dilute chlorine stream known as tail gas from the liquid chlorine plant. Some chlorine producers sell their hydrogen to merchant hydrogen producers who further purify and liquefy the chlor-alkali electrolytic hydrogen for metallurgical use and rocket fuel.

The chlorine from the cell is saturated with water at the cell temperature, and contains lesser amounts of NCl_3, organic chlorides, and salt from brine entrainment. The gas is cooled to knock out the water, and this chlorinated water is returned to the brine plant. Then the cooled gas is dried by passage through towers of sulfuric acid. Strong (98%) acid is fed to one end of the drying train at

Fig. 12.13. Bipolar and monopolar cells and circuit layouts: (a) bipolar cell electrolyzer; (b) monopolar cell electrolyzer; (c) bipolar circuit layout; (d) monopolar circuit layout.[17]

a rate equivalent to the removal of weak (65%) acid from the opposite end. The waste acid often is used to neutralize alkaline wastes, and the reconcentration and recycling of this acid presents a challenge to a chlor-alkali manufacturer.

The moisture content of the chlorine at this stage usually is low enough to allow the use of steel in the remainder of the process. The gas is compressed, usually to whatever pressure is required by the user plant, or to sufficient pressure to liquefy it, the exact pressure depending upon the refrigeration capacity and the liquid chlorine storage requirements of a given plant. Figure 12.15 is the liquefaction curve for chlorine.[28]

In nothern climates, where ambient temperatures can fall below those required to liquefy chlorine at rather modest pressures, it is necessary to provide means for maintaining the temperatures in chlorine-gas distribution lines above the liquefaction point. This can be done by tracing or wrapping the lines with heating tape, using steam lines, and insulating the pipes. These precautions are especially important in organic process plants, which

Fig. 12.14. Simplified flow sheet for chlor-alkali plant.

Fig. 12.15. Liquefaction curve for chlorine.

can run without incident while using gaseous chlorine but could explode if fed liquid chlorine.

Organic impurities and trace amounts of NCl_3 may be removed from gaseous chlorine in a simple extraction in which liquid chlorine is used as the extractant. The liquid bottoms flow from the extractor to a still in which chlorine is flashed off to recycle and the organic fraction (largely low molecular weight organic chlorides) is recovered for disposal by combustion to CO_2 and HCl, which is recovered for use.

There are many options in the liquefaction of chlorine, depending upon the storage requirements for the liquid and the system used to handle the tail gas. Low-pressure storage is recommended because of the reduced hazard in the event of a spill. However, this requires lower storage temperatures. Low temperatures also reduce the amount of chlorine in the tail gas from liquefaction. Yet lower temperature may not be necessary if a plant also deliberately produces bleach or HCl from the tail gas.

The components in the tail gas originate in the cells. At the anode, some water is oxidized to produce small amounts of oxygen. To this amount is added more oxygen from air, which may leak into the chlorine header from outside because the header is operated under a slight vacuum to prevent gas from escaping to the atmosphere. The cell gas typically contains about 97 percent Cl_2. Because the remaining 3 percent is noncondensable under the conditions in the liquid chlorine plant, it remains gaseous at the tail end of the liquefaction process, saturated with chlorine in an amount determined by the vapor pressure in that stage of the process. When air is deliberately added to the system to reduce the hydrogen content of tail gas to below 5 percent, the tail gas often increases

in volume. With hydrogen above that level, there is a great risk of explosion. The hydrogen comes from the cathode compartment of the cell, which is operated under a slight pressure to avoid sucking air into the hydrogen collection system. Ordinarily, this leads to quite small amounts of hydrogen in the cell chlorine gas. But when diaphragm problems arise or pinholes develop in the membrane, the level can rise to undesirable values. In good plant practice, this is monitored continuously.

The cell liquor flowing from the cathode compartment of each cell is collected for evaporation to 50 percent product caustic. The diaphragm cell liquor contains 10 to 12 percent NaOH, 14 to 16 percent NaCl, 0.1 to 0.5 percent Na_2SO_4, 0.02 to 0.04 percent Na_2CO_3, and 0.01 to 0.2 percent $NaClO_3$. The evaporation is carried out in reverse feed nickel-lined multi-effect systems (i.e., feed liquor enters the opposite end of the train with respect to the primary steam entry). With rapidly rising fuel costs, it has become economically justifiable in the construction of new plants to install quadruple-effect evaporators. Previously triple-effect systems were installed. Most caustic evaporation still is done in triple-effect.

Because of the high salt content of the feed liquor originating from diaphragm cells, the evaporation system is really an evaporative-crystallizer. Salt is removed from each effect in the evaporaation train. Hydroclones, centrifuges, settlers, and filters are used extensively to remove the salt as efficiently as possible and return it to the brine plant with as little caustic loss as possible. In this regard, good plant practice strikes an economic balance between caustic loss and the addition of wash water to the system, which requires additional heat input.

The sulfate content of caustic evaporator feed can cause a problem during evaporation. When the caustic concentration reaches about the 32 percent level, the triple salt $NaCl \cdot Na_2SO_4 \cdot NaOH$ crystallizes out along with NaCl crystals. To avoid sulfate buildup in the system, which would lead to excessive caustic loss, the sulfate must be purged from the system. If it is recycled with the salt, its concentration will rise to unacceptable levels. Various schemes are used, but all plants purge sulfate at some point.

The hot 50 percent caustic from the evaporator train is cooled and filtered to remove salt down to the solubility limit. Typical analysis of diaphragm cell product caustic, on a 50 percent NaOH basis (actual caustic content varies from 48.5% to 50.5%) is: 1.0 percent NaCl, 0.15 percent Na_2CO_3, 0.025 percent Na_2SO_4, 0.15 percent $NaClO_3$, 5 ppm Fe, and a maximum of 30 ppm combined heavy metals (Sb, As, Bi, Cd, Cu, Pb, Hg, Ag, and Sn). Most often the billing weight of a shipment of caustic is based on its Na_2O content, equivalent to 76 percent Na_2O:

$$\frac{(\text{pounds solution}) \times (\%\, Na_2O)}{76} = \text{net pounds NaOH } (Na_2O \text{ basis})$$

Actually, pure NaOH contains 77.5 percent Na_2O. The value 76 percent is derived from tradition, based on the fact that many decades ago the purest form of commercially available caustic contained 76 percent Na_2O. In isolated cases, where the market is far removed from the source, the caustic is further concentrated and shipped hot as 73 percent NaOH. This saves on freight costs, but the solution must be kept hot because it freezes at about 160°F. Similarly, 50 percent caustic must be maintained above 60°F.

In selected markets, such as the rayon industry, low salt levels are required (typically 100 ppm). Mercury cell and membrane cell caustic are ideally suited for these markets. Diaphragm cell caustic is purified further in a few installations, but primarily to remove $NaClO_3$, which is corrosive to the equipment used to manufacture anhydrous caustic. Chlorate alone may be removed by reaction with a reducing agent, such as sugar, which is injected into the hot caustic feed to the anhydrous concentrator system. This has the disadvantage of increasing both the salt and the carbonate levels in the caustic.

To remove chlorate and other salts, the 50 percent diaphragm cell caustic may be extracted with liquid ammonia in a pressurized system.[29] The ammonia fraction then is processed through a stripper to remove and recycle the ammonia. The alkaline stripper bottoms are useful in neutralizing acidic waste streams. The purified caustic is evaporated further, and then fed to an anhydrous concentrator, typically an Inconel falling film evaporator, heated with molten salt. The anhydrous caustic, containing at least 97.5 percent NaOH, is marketed as one solid mass in drums, as flake caustic, or more desirably as beads or prills. The last are marketed in bulk or bags.

Since the binary system H_2O–NaOH contains several hydrates that may be conveniently crystallized, another possibility exists for purifying caustic. Of special interest are the dihydrate, $NaOH \cdot 2H_2O$, which melts at about 55°F, and $NaOH \cdot 3.5H_2O$, which melts at 60°F. The impurities concentrate in the mother liquor; purified caustic then is produced by melting the crystals. The most recent attention to this method of purification has been in Japan.[30] These processes are more difficult than they appear, because of the high viscosity of concentrated caustic solutions near their freezing points and the resultant challenges of both adequately transferring heat and separating the crystals from the mother liquor. But with the advent of the membrane cell, there may be more development activity in this area.

In the handling and storage of 50 percent caustic, it is important to limit the amount of iron pickup. The iron spec is 5 ppm on a solution basis. Stainless steel offers little advantage over mild steel with respect to iron contamination. Furthermore, hot, salt-containing caustic leads to severe stress corrosion cracking of stainless steel. Carbon steel storage tanks often are lined with an epoxy-phenolic or neoprene latex coating. Iron pickup from transfer piping increases with increasing temperature and increasing linear velocity of pumping. Unnecessary recirculation or pumping through steel lines should be avoided.[31] Iron, copper, and nickel impurities are especially undesirable for the bleach manufacturer.[32]

About 75 percent of the caustic produced is concentrated. The remainder is used directly as alkaline cell liquor—as, for example, in the conversion of propylene to propylene oxide by the chlorhydrin process. Similarly, there is some chlorine produced by methods that do not produce caustic, as shown in Table 12.18. Fused chloride salt electrolysis produces chlorine in the manufacture of magnesium metal by the Dow process, and of sodium metal in the Downs cell. The only other process of note is the Kel-Chlor process.[33] This process converts by-product HCl to chlorine by oxidation with NO_2 through the intermediates NOCl and $HNSO_5$.

A small, but important, variation of NaCl electrolysis substitutes KCl as the feed. Both mercury cells and diaphragm cells are used, producing chlorine and KOH (caustic potash). The KOH is concentrated for sale as a 45 percent solution or in solid form, containing 88 to 92 percent KOH. A big use for KOH is in the manufacture of liquid soaps and detergents. The analogous sodium soaps and detergents are generally solids.

Table 12.19 gives the 1990 U.S. prices for chlorine and for caustic soda in its various forms.[6] These are typical values. For significant volumes, much movement of chlor-alkali goods is contracted below market price.

The entire chlor-alkali world must constantly perform a balancing act. The electrolysis of salt brine yields almost equal amounts of chlorine and caustic. It is rare that the growth or use-rates for the two products are equal. Furthermore, the chlor-alkali balance problem extends to related products such as soda ash,

TABLE 12.18 U.S. Chlorine Production by Process

Process	*Percent*
Diaphragm cells	75.4
Mercury cells	16.7
Membrane cells	4.9
Fused salt electrolysis	1.2
Other	1.8
Total	100.0

TABLE 12.19 1990 Value of Chlor-Alkali Products[6]

Product	*$/Ton (100% Basis)*
Liquid chlorine	125
50% caustic	300
50% caustic (rayon)	330
73% caustic	350
Solid caustic	560
Flake caustic	560
Bead caustic	560

lime, salt cake, and hydrochloric acid, even to the extent of influencing (and being influenced by) the current economics of chlorinated hydrocarbons.

For example, in the 1960s the demand for chlorine was so much greater than for caustic that concerted efforts had to be made to move caustic into new markets, such as the glass industry. Another effect of the same economic climate was a glut of HCl from the tremendous growth in vinyl chloride and a host of other chlorinated organic derivatives. Necessity being the mother of invention, this situation led to the development of the oxy-chlorination process for the production of vinyl chloride, resulting in no net production of HCl; to a route for producing methyl chloride using HCl and methanol; to the Kel-Chlor process for production of chlorine without caustic; and to the recycle of by-product HCl to chlorine by electrolysis.[34-36] There was a host of developmental activity in both the United States and Europe to convert caustic to soda ash and to improve the old Deacon process, which is a nonelectrolytic, catalytic conversion of HCl to chlorine.[37] In another case, a large consumer of sulfuric acid, the steel industry, was converted to the use of HCl for pickling steel. Huge amounts of caustic cell effluent were used to replace lime in a variety of applications.

By the mid-1970s, the picture was changing. Environmental pressures on fluorocarbon derivatives (derived from chlorinated compounds and creating by-product HCl), and on some of the chlorinated solvents, impacted negatively on chlorine growth. During the late 1980s, the pulp and paper industry started to come under intense environmental scrutiny due to the presence of total organic halides (AOX) in effluents from bleaching operations in kraft mills. As a result, it appears that the industry will voluntarily reduce chlorine usage by 50 to 70 percent in the early 1990s and use chlorine dioxide in its place. In order to reduce the AOX discharges from over 3.0 kg/ton of pulp manufactured to less than 1.5 kg/ton of pulp, the pulp and paper industry also will be pushed to use a more expensive oxygen delignification technology that would lead to further reductions in chlorine use in pulp bleaching applications. The result of this would be zero or negative growth for chlorine, a situation that has tremendous implications for caustic availability.

The swings in the chlor-alkali industry during the 1980s, coupled with environmental pressures on the use of mercury and asbestos in chlor-alkali cells, led to numerous plant closures. In 1989 there were 26 U.S. companies producing chlorine at 51 sites with an operating capacity of 12,275 short tons. Dow Chemical U.S.A. accounted for almost a third of U.S. capacity. The top four producers shown in Table 12.20 accounted for over 70 percent of the installed capacity.

In early 1992, Dow announced it would shut down all chlorine producing and using facilities at its Sarnia, Canada, plant by mid-1993. Essentially this would decrease North American chlorine capacity by about 1500 tons per day.

The U.S. chlorine supply–demand as projected by Dow analysts[39] at the end of 1990 is shown in Fig. 12.16. Dealing with multiple issues on the environmental front and the possibility of a recession, the chlor-

TABLE 12.20 Major North American Chlorine Producers[6]

Dow Chemical U.S.A.	30.7
Occidental Petroleum Corporation	24.4
PPG Industries, Inc.	11.6
Olin Corporation	7.2
22 other companies	26.1
	100.0

Fig. 12.16. U.S. chlorine supply and demand, based on many market variables. The more optimistic forecast assumed the pulp and paper segment would proceed only to a 50% reduction level, resulting in a 0.9% average annual growth rate for the 1988 to 1995 period. The second outlook assumed the recession forecast for the economy and a 70% reduction in demand from the pulp and paper segment, which would lead to no growth for chlorine. These two outlooks represent extreme cases. Dow analysis predicted that market growth figures would fall somewhere between the two. (*Courtesy the Dow Chemical Company.*)

alkali industry was faced with many challenges entering the 1990s.

The continued growth of vinyl chloride monomer, both domestically and globally, will ensure a sink for the bulk of the chlorine and adequate though tight caustic supplies for the pulp and paper industry. The tight caustic soda market will permit caustic producers to recover a larger portion of their chlor-alkali production costs from merchant caustic sales. The portion chargeable to chlorine will be lowered, and more ethylene dichloride capacity for VCM/PVC production will be installed.

Sodium carbonate will continue to make inroads into the caustic market. In some cases, even lime or magnesium hydroxide will be substituted for caustic. Consideration will be given to producing caustic without coproduced chlorine. The old lime soda process for the conversion of lime and soda ash into caustic will be practiced. There may even be a process to produce sodium hydroxide from sodium carbonate electrolysis, although there is currently no commercial production via this route.

HYDROCHLORIC ACID

The predominant source of HCl is as a by-product from the chlorination of hydrocarbons, including the incineration of chlorination organic wastes in thermal oxidizers. Hydrogen and chlorine can be burned together to produce HCl. Hydrochloric acid produced by the reaction of salt and sulfuric acid is not economically competitive in North America. By 1990, more than 93 percent of the HCl was being produced as a coproduct from the chlorination of hydrocarbons. HCl burners to convert chlorine plant tail gas to HCl accounted for 5 percent of total HCl production, and less than 2 percent of the HCl was produced by the salt and sulfuric acid process. It has been estimated that by 1990 over 6 million tons of HCl was being produced in North America. Accurate data for hydrochloric acid are difficult to obtain because of the large number of producing and consuming companies, the variety of grades, and the general reluctance of many producing companies to discuss HCl production and use. The production of HCl depends on demand for the primary product (vinyl chloride monomer, isocyanates, fluorocarbons, etc.). Current prices for HCl ranges from $300 to $340 per ton for the anhydrous product.

Hydrogen chloride is marketed both as anhydrous HCl and as hydrochloric acid, usually a 32 percent solution. The absorption of HCl in water customarily is achieved in a falling film absorber, followed by a packed tails tower that vents inerts to the atmosphere. The weak acid from the secondary absorber is fed to the primary absorber. The degree of absorption of the HCl is dependent on the

concentration of inerts in the gas stream. The heat of absorption of HCl in water is about 700 Btu/lb of HCl, which requires that the primary absorber be water-cooled.

Acid made by burning hydrogen in chlorine tail gas is quite pure and requires no further treatment. By-product acid requires purification in packed scrubbers in which the scrubbing medium is the hydrocarbon, which is circulated in the scrubbing tower before being sent to the chlorinator. A refrigerated cooling system frequently is employed to maintain a low temperature in the circulating scrubber liquor in order to maximize absorption efficiency.

BROMINE AND BRINE CHEMICALS*

Although bromine once was recovered from seawater in the United States, and there still are seawater operations in France and Spain, the most important source of bromine is brines and bitterns (the mother liquor from a saltworks). Seawater contains only 65 ppm of bromine, so very large volumes of it must be processed to recover bromine. Major sources of commercial bromine are underground brines in Arkansas (which contain 3000 to 5000 ppm bromine), China, Russia, and the United Kingdom. Bitterns from mined potash in France and Germany, seawater bitterns in India, Italy, and Japan, and bitterns of potash production (which contain 12,000 ppm bromine) from Dead Sea brines in Israel are the other sources.[40] The United States now is the largest producer, with most U.S. production in the area around Magnolia, Arkansas. In 1990, U.S. bromine capacity was 575 million lb/year. Israel is the second largest producer, and its production is growing by far the fastest. The Former Soviet Union is third in production.

There are four major steps in bromine production: (1) oxidation of bromide to bromine with chlorine; (2) stripping of bromine from the aqueous solution; (3) separation of bromine from the vapor; and (4) purification of the bromine. The two major stripping vapors are steam and air. Steam is used when the concentration of bromine in brine is greater than 1000 ppm. The advantage of using steam is that bromine can be condensed directly from it. Air is used when seawater is the source of bromine because very large volumes of stripping gas are needed, and steam would be too expensive. When air is used, the bromine needs to be trapped in an alkaline or reducing solution to concentrate it.

In a newer steam-based process, vacuum is used, which, by matching the vapor pressure of the hot brine, causes bromine to be stripped out without the need for steam to heat the brine. Claimed benefits are increased plant capacity, reduced chlorine loss, less lime needed to treat the spent brine, lower plant maintenance costs, and decreased atmospheric emissions.[41,42]

Table 12.21[6] gives the market distribution for bromine, which in 1990 was selling for 43 to 46 cents/lb. The largest use of bromine is in the production of flame retardants. Additive-type flame retardants, such as decabromobiphenyl oxide, are physically blended into polymers; reactive types, such as tetrabromobisphenol A, chemically react during polymer formation. Brominated polymers are used in flame-retardant applications. Other large-volume bromine chemicals are methyl bromide, used as an agricultural fumigant, and ethylene dibromide, used in lead antiknock fluids outside the United States. Ethylene dibromide was once the largest-volume

TABLE 12.21 Market Distribution of Bromine[6]

Market	*Percent*
Flame retardants	27
Methyl bromide	25
Ethylene dibromide	20
Drilling fluids	10
Water treatment chemicals	
Other (i.e., photographic, rubber additives, etc.)	14
Exports	5

*The material in this section was contributed by Philip F. Jackish, Ethyl Corp., Baton Rouge, LA, and John Pavlick, Dow Chemical U.S.A., Ludington, MI.

bromine chemical, but the phase-out of lead in gasoline and its banning as an agricultural fumigant have reduced its markets. Other bromine compounds find use as pesticides and as intermediates in various chemical processes.

A fairly new major use for bromine is in clear completion fluids used in crude oil production. These are concentrated solutions of $CaBr_2$ or $CaBr_2$–$ZnBr_2$, which have a high density. Several hundred pharmaceuticals contain bromine atoms. Other moderate-sized markets for bromine compounds are in photography, water treatment chemicals, extinguisher fluids, and dyes.

Miscellaneous bromine uses are in catalysts, fluxes, precious metal recovery, hair care products, food additives, flotation agents in ore treatment, solvents, refrigerants, quartz-halide light bulbs, some lasers, some photovoltaic batteries, and some electrically conductive polymers. Zinc-bromine storage batteries are being developed for use as load-leveling devices in electric utilities,[43] and this could became a significant market for bromine.

In Michigan, this debrominated brine (containing 10–20% $CaCl_2$ and 3–10% $MgCl_2$) is processed further to produce calcium chloride and magnesium hydroxide, as shown in Fig. 12.17.

Magnesium hydroxide is produced by precipitation of the magnesium ions with $Ca(OH)_2$ obtained from dolomitic limestone. The limestone is calcined in a rotary kiln to produce lime; and the lime is reacted with water to give $Ca(OH)_2$, which reacts with the Mg ions in the brine to give $Mg(OH)_2$. One-half of the $Mg(OH)_2$ comes from the dolomitic lime. The slurry of $Mg(OH)_2$ in $CaCl_2$ brine is thickened in settling tanks and filtered, and the $CaCl_2$ is washed out of the cake with water.

The major use of magnesium hydroxide is as the raw material for magnesium oxide

Fig. 12.17. Brine chemicals in Michigan. (*Courtesy Dow Chemical Company.*)

TABLE 12.22 Market Distribution of Calcium Chloride[6]

Market	*Percent*
Pavement deicing	35
Dust control/road stabilization	22
Industrial processing	21
Oil and gas well fluids	10
Concrete	5
Miscellaneous	4
Exports	3

production, which is used in refractory brick. The hydroxide is converted to the oxide by heat treatment in a rotary kiln or multiple hearth furnace. Other uses are in pulp and paper, Epsom salts, waste water treatment, flame retardants, and smoke supressants.

The effluent from the magnesium hydroxide process is a 25 percent calcium chloride brine. The water in the brine is boiled away with steam in triple-effect vacuum evaporators to make 32 to 45 percent $CaCl_2$. Alkali chlorides precipitate in the evaporators and are removed by settlers and centrifuges. Additional water is removed in evaporators and dryers to attain anhydrous calcium chloride at concentrations from 78 to 94 percent.

Calcium chloride is both hygroscopic and deliquescent. Thus, solid material will absorb moisture from the air or snow and ice until it dissolves, and the solution will continue to absorb moisture until an equilibrium is reached between the vapor pressure of the solution and that of the air. At 95 percent relative humidity, one pound of anhydrous calcium chloride will absorb 17.3 lb of water. The ability of calcium chloride to absorb moisture makes it useful in many construction, commercial, and industrial applications. Table 12.22 gives the market distribution of calcium chloride.[6] The 1990 calcium chloride price for 78 percent flakes was about $215/ton.

BLEACHES

Since the fledgling days of industrial chemical practice, the bleaching properties of chlorine have been in demand. It is noteworthy that the beginnings of the industry, some two hundred years ago, centered upon materials and methods of bleach manufacture. Much of modern chemical engineering practice has derived from the struggles and developments of this heritage, particularly in the closing years of the last century, when the electrolytic production of chlorine and caustic began to take hold. Rampant spread of the great killer diseases such as cholera and typhoid was shown to be easily eliminated by the treatment of wastewater and drinking water with small amounts of bleach. Textile manufacturers and the paper industry had established the value of bleach in their operations early on.

Prior to the development of the liquid chlorine industry, the only means of transporting chlorine was as bleach. Common practice in the early years was to absorb the chlorine in hydrated lime, $Ca(OH)_2$, to form chloride of lime, $CaOCl_2$. Easily transported as a solid, this chemical contains the equivalent of 35 percent Cl_2. When dissolved in water, it forms equal moles of $CaCl_2$ and $Ca(OCl)_2$:

$$2CaOCl_2 \text{ (dissolved)} \rightarrow CaCl_2 + Ca(OCl)_2$$

It is not the same as calcium hypochlorite, which contains over 99 percent available chlorine. (The term "available chlorine" refers to the oxidizing power of a bleach equivalent to that of free chlorine, as determined by the standard acidic-KI-thiosulfate titration.)

The most practical route to achieving a high available-chlorine content in a lime-based bleach involves chlorinating a slurry of lime in caustic,[8] followed by cooling to $-10°F$. Crystals are centrifuged off and added to a slurry of chlorinated lime containing $CaCl_2$ in an amount equivalent to the NaOCl content of the original crystals. Warming this solution produces crystals of $Ca(OCl)_2 \cdot 2H_2O$, containing around 3 percent lime as an impurity. This material, which is granulated, dried, and packaged for sale, has an available chlorine content over 70 percent.

Another approach uses chlorine monoxide and water to form a solution of HOCl, which is neutralized with a lime slurry, spray-dried, and granulated to yield a product containing 70 percent available chlorine.[32]

The lime-based products have the disadvantage of containing insoluble components, which has led to the necessity for settling and sludge disposal in some applications. These components are especially objectionable in laundry and dishwashing uses. A granular bleach based on LiOCl[30] is quite useful in these applications because it is completely soluble. It is prepared as a 35 percent solution of LiOCl by mixing solutions of LiCl and NaOCl, from which NaCl precipitates. The solution is evaporated, and the solid (containing LiOCl and NaCl) is dried and formulated with salt cake to a white, free-flowing, granular product that retains its bleaching power (35% available chlorine) reasonably well during storage.

In the bleaching of fibers for the manufacture of paper and textiles, sodium chlorite ($NaClO_2$) and/or chlorine dioxide (ClO_2) often is used because of the ability to achieve excellent whiteness without as much fiber degradation as occurs in chlorine or hypochlorite bleaching. For example, in the production of kraft paper, known for its strength, the product is brown when conventional bleach is used because bleaching to whiteness would destroy its strength. By using ClO_2, a high-strength, white product is obtained.

An additional benefit of bleaching with ClO_2 is that the waste lignin degradation products generally present fewer disposal problems than do waste products of chlorine-based bleaches.[44] Kraft pulp bleaching sequences that involved chlorine in the first stage (C), followed by caustic extraction (E), followed by two stages of hypochlorite bleaching (H), are being changed to vary the chlorine/chlorine dioxide ratio (D) in the first stage, followed by two stages of chlorine dioxide (i.e., CEHEH replaced by C_DEDED). Regulations are being contemplated regarding maximum discharge levels of chlorinated organics and dioxins. Sodium hypochlorite has been linked to the formation of harmful chloroform emissions from bleach plants. It is expected that chlorine dioxide use in bleaching will increase, both as a partial substitute for elemental chlorine and as a replacement for sodium hypochlorite. It has been shown that chlorine dioxide substitution in the first-stage bleaching sequence above the 50 percent level can reduce total organic halide (AOX) discharge by up to 50 percent and eliminate the formation of dioxins.[45] By the mid-1990s the use of oxygen and hydrogen peroxide in the lignin extraction stage (E-stage) is expected to be common, and it eventually could dampen ClO_2 demand. Also ozone bleaching may limit growth in ClO_2 bleaching.

The generation of ClO_2 from chlorate involves a complex reduction of the chlorine atom of chlorate from the +5 oxidation state to the +4 oxidation state in an acid medium. Chlorine dioxide is a very unstable molecule. It cannot be economically transported, but is easily produced in situ. Treatment of a solution of $NaClO_3$ and NaCl with acid (H_2SO_4 or HCl) produces ClO_2, which is immediately absorbed for use in bleaching. The ClO_2 generator must be operated with care to avoid ClO_2 concentrations above 10 percent, which can lead to explosion from self-decomposition. The molar ratio of ClO_2 to Cl_2 formed in the generator is 2:1. If this level of Cl_2 is objectionable, the gas mixture from the generator, which is mostly air (for safety), is passed through a tower through which chilled water is circulated. This dissolves the ClO_2, but only 25 percent of the Cl_2; so the remaining Cl_2 is scrubbed with alkali. This treatment raises the molar ratio to $8ClO_2:1Cl_2$.

Over the years a number of commercial chlorine dioxide generators have been developed. These generators can be classified according to the type of acid media (i.e., sulfuric acid or hydrochloric acid) and the type of reducing agent (i.e., sulfur dioxide, methanol, sodium chloride, or hydrochloric acid). In most processes, some elemental chlorine is liberated as a by-product. The generator using sulfuric acid produces sodium sulfate and/or waste sulfuric acid, which must be disposed. A new process (R-8 generator[46]) has been commercialized that is capable of eliminating by-product chlorine (i.e., so that there is no hypochlorite) and produces reduced quantities of by-product salt cake.

Although chlorine dioxide has replaced sodium hypochlorite use in kraft pulp bleaching, hypochlorite bleach is widely used as an ordinary household bleach because of its ease of manufacture, cost, and handling convenience. Sodium hypochlorite is produced mostly by reaction of liquid chlorine with 20 percent NaOH. Soda bleaches vary in their NaOCl content up to about 15 percent; and the higher the NaOCl content is, the higher the required excess of caustic to maintain stability. For this reason, household bleach contains 5.5 percent NaOCl. In the trade it is common to speak of the available chlorine level of soda bleach as the "trade percent." This is related to the NaOCl concentration, expressed in grams per liter, as:

$$(\text{Trade \% available } Cl_2) = (\text{GPL NaOCl}) \times 0.0953$$

$$(\text{Weight \% available } Cl_2) = (\text{Trade \%})/(\text{Specific gravity})$$

Household bleach has a weight % available Cl_2 of 5.2, and a trade % of 5.6, which is the percent of available chlorine per volume of solution. It is to be noted that not only is sodium hypochlorite an oxidant, but it yields oxygen on decomposition:

$$2NaOCl \rightarrow O_2 + 2NaCl$$

SODIUM CHLORATE

The rapidly growing use of ClO_2 in the pulp and paper industry has led to the rapid growth of sodium chlorate, $NaClO_3$, production in recent years. Sodium chlorate is rapidly produced in solution form by the electrolysis of NaCl brine in a cell that is very similar to a diaphragm chlor-alkali cell, except that it has no diaphragm. The overall reaction is as follows:

$$NaCl + 3H_2O \rightarrow NaClO_3 + 3H_2$$

In this process the chlorine and caustic produced in the cell react immediately, and the cell liquor is kept hot in an additional vessel that provides sufficient residence time for all the hypochlorite to disproportionate into chloride and chlorate:

$$3NaOCl \rightarrow NaClO_3 + 2NaCl$$

For optimum operation, the cell temperature is maintained at 65 to 75°C with a pH of 6.9. Current efficiencies of over 90 percent have been reported with a power consumption of 4500 to 5800 kWh/metric ton of sodium chlorate produced.

Various cell designs are offered by licensers of sodium chlorate technology, the anode design being the most significant variable in the chlorate cell. Coated titanium metal anodes and steel cathodes are used, and small amounts of dichromate are added to the cell liquor to passivate the steel so that corrosion is minimized. The cell liquor contains about 50 percent $NaClO_3$, which may be used as is, or evaporated to 75 percent to drop out salt, filtered, and cooled to crystallize out $NaClO_3$.

In 1990, the total North American sodium chlorate capacity was 1.1 million tons/year with 67 percent of this capacity in Canada because of its lower energy costs. Over 94 percent of the North American sodium chlorate is used for the manufacture of on-site ClO_2 for wood pulp bleaching, with the remainder being used for the manufacture of other chlorates and in perchlorates, chlorites, herbicides, and water treatment and mining. The 1990 price for bulk sodium chlorate was about \$420/ton for the crystalline product. As a result of increasing demand for chlorine dioxide in pulp bleaching, a number of North American sodium chlorate producers have annomic plant expansions, particularly in Canada because of lower electric power costs. By 1995, once pulp mills reach their final chlorine dioxide substitution levels, sodium chlorate growth is expected to drop.

Akin to $NaClO_3$ is the similar manufacture of bromates, starting with NaBr, or KBr, cell feed. Bromates also are produced by bromination of the respective carbonates. The bromate business is extremely small by comparison to other oxidants, but bromates are important in the bread-making industry for maturing flour and in conditioning dough. They also

are used in permanent wave lotions and in the manufacture of dyes. The bromates usually are sold as specialty chemicals. Both chlorates and bromates have substantial oxidizing power, and care must be taken in storage and handling of these solids. Contact with reducing agents and combustibles can be diastrous.

REFERENCES

1. Kauffman, D. W., *Sodium Chloride*, Van Nostrand Reinhold, New York, 1960.
2. Leddy, J. J., *J. Chem. Ed.*, **47**, 386 (1970).
3. *Chemical Economics Handbook*, Stanford Research Institute, Menlo Park, CA, 1988.
4. *Chemical Economics Handbook*, Stanford Research Institute, Menlo Park, CA, 1990.
5. *Chemical Economics Handbook*, Stanford Research Institute, Menlo Park, CA, 1987.
6. *Chemical Marketing Reporter*, Schnell Publishing Co., New York (Oct. 1990).
7. "Industrial Mineral Background Paper 2," Law, Sigurdson and Associates, Toronto, Ontario and SRI International, Menlo Park, CA.
8. Sconce, J., "Chlorine," ACS Monograph No. 154, p. 180 (1962).
9. U.S. Patent 4,793,906.
10. Van Wazer, J. R., *Phosphorous and Its Compounds*, Interscience, New York, 1958.
11. U.S. Patent 1,385,595.
12. U.S. Patent 1,132,640.
13. German Patent 249,222.
14. Leddy, J. J., *J. Chem. Ed.*, **57**, 640 (1980).
15. *Chemical Week*, p. 36 (Apr. 12, 1989).
16. Leddy, J. J., et al., *Kirk-Othmer Encyclopedia of Chemical Technology*, John Wiley, New York, 1978.
17. Burney, H. S., and White, R. E., Dow Chemical Company, presented at AIChE Symposium: Tutorials in Electrochemical Engineering and Technology—IIIh, Sept. 1989.
18. Mantell, C. L., *Electrochemical Engineering*, 4th ed., McGraw-Hill, New York, 1960.
19. Kuhn, A., *Industrial Electrochemical Processes*, Elsevier Publishing Co., 1971.
20. U.S. Patent 3,718,627.
21. Grot, E. G., *Chem.-Ing. Tech.*, **44**, 167 (1972).
22. Seko, M., New York A.C.S. Meeting, Apr. 1976.
23. Beer, H., *Chem. Tech.*, **9**, 150 (1979).
24. U.S. Patent 3,632,498.
25. German Patent 71937.
26. Latimer, W. M., *Oxidation Potentials*, Prentice-Hall, Englewood Cliffs, NJ, 1952.
27. Dotson, R. L., *Chem. Eng.*, p. 106 (July 17, 1978).
28. Kapoor, R., and Martin, J. J., "Thermodynamic Properties of Chlorine," Eng. Res. Inst., University of Michigan (1957).
29. *Chem. Met. Eng.*, **51**, 119 (1944).
30. U.S. Patent 3,983,215.
31. Considine, D. M., *Chem. Proc. Tech. Encyclopedia*, p. 233 McGraw-Hill, New York, 1974.
32. White, G. C., *Handbook of Chlorination*, Van Nostrand Reinhold, New York, 1972.
33. Van Dyk, C. P., *Chem. Eng. Progress*, **69**, 47 (1973).
34. U.S. Patent 3,635,804.
35. Grosselfinger, F. C., *Chem. Eng.* (Sept. 14, 1964).
36. U.S. Patent 3,242,065.
37. U.S. Patent 3,210,158.
38. U.S. Patent 3,989,807.
39. Newsletter for Dow's Chlor-Alkali Business, Summer 1989.
40. Lyday, P. A., "Bromine," in *Minerals Yearbook 1987*, Vol. 1, p. 172, Bureau of Mines, Washington, D.C.
41. U.S. Patent 4,719,096.
42. U.S. Patent 4,725,425.
43. Veziroglu, T. N., (Ed.), "Alternate Energy Sources," *Proc.* Miami Int. Conf., 6th Meeting, 1983, Vol. 1, pp. 327–333, Hemisphere, Washington, DC.
44. Reeve, D. W., and Earl, P. F., *Pulp and Paper Canada*, **90**, 4 (1989).
45. Fleming, B. I., et al., *Pulp and Paper Canada*, **89**, 12 (1988).
46. Krauza, K. E., Canadian Pulp and Paper Association, Spring Meeting, 1989.

13

Industrial Gases

Glenn W. Mattern*,**

OVERVIEW

Industrial gases are considered to be those gases used in their pure form by industry in large quantities. They are used in industry as gases, liquids, or cryogenic liquids.

The gases we consider to be industrial gases have been in use for many years. For example, nitrous oxide was used as an anesthetic as early as 1799. By 1608, carbon dioxide had been identified as a specific substance acquired by several means. Some gases for industrial use are produced by means of cryogenics, which is the science and technology of dealing with the behavior of materials at extremely low temperatures.

In this chapter three types of gases are discussed: oxidizers, inert gases, and flammable gases. The gases that fall into the category of inert are nitrogen, argon, helium, and carbon dioxide. The oxidizers are oxygen and nitrous oxide. Acetylene, liquefied natural gas (LNG), and hydrogen are the flammable gases. These elements and compounds by no means encompass all gases considered to be industrial gases, but they represent the major gases used and produced in industry. (See Table 13.1.)

Nitrogen, oxygen, and argon produced for industry are primarily manufactured by the cryogenic air separation process, in which air is liquefied and separated into its major constituents (oxygen, nitrogen, argon) by the use of cryogenic technology. Table 13.2 gives the cryogenic boiling point temperatures and concentrations for some of the gases present in the atmosphere.

The cryogenic air separation process starts with air entering the plant through air filters, where it is compressed and cooled. The air is passed through heat exchangers for further cooling and for removal of water vapor and carbon dioxide by freezing. Alternatively, a molecular sieve can be used to remove the water and carbon dioxide by adsorption before separation of the air. When the air reaches a temperature of $-170°C$, it is sufficiently free of water and carbon dioxide and ready to be passed into the distillation column.[4]

Separation occurs by distillation. (See below, Fig. 13.5.) The distillation method used is

*Air Products and Chemicals, Inc.
**The author wishes to acknowledge the assistance of Katanya Bolling in preparing this chapter.

TABLE 13.1 Properties of Industrial Gases

Gas Property		*Helium-4* *He*	*Para-Hydrogen* *p-H_2*	*Normal-Hydrogen* *n-H_2*	*Nitrogen* *N_1*	*Air*	*Argon* *Ar*	*Oxygen* *O_1*	*Methane* *CH_4*	*Acetylene* *C_2H_2*	*Carbon Dioxide* *CO_2*
(Basis)		(1)(2)	(1)(3)	(1)(3)(4)	(1)(5)	(6)	(1)(7)	(1)(8)(9)	(1)(3)	(10)(11)	(1)(8)(11)
Atomic/Molecular Weight		4.0026	2.0159	2.0159	28.0134	29.964	39.948	31.999	16.042	26.037	44.009
Normal Boiling Point	(°F)	−452.07	−423.19	−422.97	−320.42	−317.66 to −312.54 (B)	−302.56	−297.33	−258.75	−118.5 (C)	−109.33 (C)
Triple Point	(°F)	−455.75 (A)	−434.83	−434.55	−346.00	—	−308.83	−361.84	−296.45	−115.3	−69.90
	(PSIA)	0.730 (A)	1.021	1.045	1.818	—	9.994	0.02201	1.703	17.55	5.65
Critical Point	(°F)	−450.31	−400.31	−399.93	−232.44	−221.33 (B)	−188.05	−181.39	−116.59	96.5	87.8
	(PSIA)	32.99	187.51	187.51	493.00	545.01 (B)	711.52	731.42	666.88	905.4	1071.6
Gas Density											
@ NTP	(lb/CF)	0.01033	0.0052091	0.0052091	0.072456	0.074911	0.10336	0.082786	0.041553	0.067814	0.11438
@ STP	(lb/CF)	0.011150	0.0056113	0.0056113	0.078043	0.080721	0.11137	0.089209	0.044789	0.073184	0.12341
Vap. Density											
@ NBP	(lb/CF)	1.0543	0.083529	0.083529	0.2879	0.28057 (B)	0.3606	0.2795	0.1134		0.1462
Liq. Density											
@ NBP	(lb/CF)	7.8019	4.4190	4.4190	50.480	54.617	87.023	71.231	26.379	38.73 (D)	97.51 (D)
Specific Heat											
@ NTP	(BTU/lb-°F)	1.241	3.558	3.415	0.2487	0.2404	0.1247	0.2197	0.1134	0.3938	0.2025
Specific Heat Ratio @ NTP		1.6667	1.3834	1.4066	1.4019	1.3995	1.6693	1.3976	1.3077	1.240	1.2871
Latent Heat											
(D) @ NBP	(BTU/lb)	8.784	191.66	191.84	85.446	87.988 (B)	69.411	91.402	219.56	354.9 (B)	246.6 (B)

NTP = 14.696 PSIA and 70°F
STP = 14.696 PSIA and 32°F

Notes: (A) lower lambda point
(B) without CO_2
(C) sublimation temperature
(D) solid density
(E) latent heat of sublimation

Basis: (1)–NBS Computer Program MIPROPS, 32-Term Pure Component Equation Of State, APCI CAPP Ithermo 18.
(2)–NBS Technical Note 631, "Thermophysical Properties Of Helium-4 ...," Nov. 1972.
(3)–NBS Monograph 168, "Selected Properties Of Hydrogen (Engineering Design Data)," Feb. 1981.
(4)–NASA SP-3089, "Hydrogen Technological Survey—Thermophysical Properties," 1975.
(5)–NBS Technical Note 648, "Thermophysical Properties Of Nitrogen ...," Dec. 1973.
(6)–NBS Extended Corresponding States Equation Of State, APCI CAPP Ithermo 19, Dry Air Composition (Mole %): 78.088 N_2, 20.948 O_2, 0.929 Ar, 0.035 CO_2.
(7)–NBS Technical Note 27, "Thermodynamic Properties Of Argon ...," Mar. 1969.
(8)–NASA SP-3090, "ASRDI Oxygen Technology Survey, Vol. IX: Oxygen Systems Engineering Review," 1975.
(9)–NASA SP-3071, "ASRDI Oxygen Technology Survey, Vol. I: Thermophysical Properties," 1972.
(10)–Local Composition Peng-Robinson Equation Of State, APCI CAPP Ithermo 17.
(11)–Average of APCI, Linde, Matheson, and MSDS published thermophysical data.

TABLE 13.2 Cryogenic Gases: Boiling Point and Concentration in Air

	Normal Boiling Point (°F)	*Dry Air Concentration (by Vol.)*
Argon	−302.6	0.93%
Nitrogen	−320.5	78.1%
Oxygen	−297.3	20.94%
Helium	−452.1	0.0005%
Hydrogen	−422.99	0.00005%
Neon	−410.7	0.0018%
Krypton	−244.0	0.00011%
Xenon	−163.0	0.000008%

referred to as "double column," as it contains two separate distillation columns operated at two different pressures. The cooled air is fed to the bottom of the high-pressure column. Air is distilled in this column where the higher boiling temperature fluid oxygen stays at the bottom and the lower boiling temperature nitrogen boils to the top as a distillate product. The crude liquid oxygen from the high-pressure column is flashed into the low-pressure column as the two columns are thermally linked by a common reboiler/condenser. Vapor is withdrawn from the middle of the low-pressure column and is passed as feed to the crude argon column. The vapor in that column is condensed and taken off as liquid crude argon. Pure nitrogen vapor is extracted from the top of the low-pressure distillation column, and liquid oxygen is taken from the bottom of the column. Figure 13.1 shows an air separation plant with a molecular sieve front-end cleanup and liquid storage tanks.

There are other means by which the atmospheric industrial gases can be produced. In addition to cryogenic air separation, there are other, non-cryogenic-based processes,

Fig. 13.1. Air separation plant including cryogenic storage tanks and cold box. (*Courtesy Air Products and Chemicals, Inc.*)

including the membrane separation process and adsorption processes. The adsorption processes include pressure swing adsorption (PSA) and vacuum swing adsorption (VSA).

The membrane separation process is used to produce a number of gases for industrial use. The membrane may consist of a microporous substructure of cellulose acetate and a thin layer of dense cellulose acetate (active layer) on the upper surface, resembling a sponge covered by a plastic film. The membrane is cast into a supporting cloth for added mechanical strength. The active layer serves as the separating barrier and, because of its thinness, provides very high transport rates. The membrane selectively allows the smaller molecular-sized components to pass through the membrane substructure, thereby accomplishing the separation process.

The VSA system is a noncryogenic process that makes use of pressure that is less than atmospheric (vacuum) and adsorbents to separate nitrogen, and carbon dioxide from air. The PSA system is a similar process for producing a particular gas by passing a pressurized gas stream through a set of adsorbent beds. The impurities are selectively adsorbed using a molecular sieve to trap the undesired components. In the production of oxygen using the VSA process, the VSA system separates air at ambient temperature by passing it through a column of adsorbent. The adsorbent, which is an inorganic crystal material, selectively adsorbs nitrogen molecules on its surface, while the oxygen molecules pass on through the adsorbent column. The final oxygen product exits the adsorber column. The nitrogen adsorbed is released when the adsorber column is reactivated by use of differential pressure from a vacuum. Once the adsorber column is reactivated, the cycle is ready to be repeated.

In the PSA system, compressed air is passed through a column of chemical adsorbent, which is a carbon-based material that essentially traps the oxygen molecules and allows the nitrogen molecule to pass through to the use point. Figure 13.2 shows the process flow of a PSA system, which can produce 98 to 99.5 percent pure nitrogen without the need of downstream purification. These adsorption systems usually are located at the customer sites. See Fig. 13.3 for an example of an on-site PSA.

Most of the industrial gases are filled in high-pressure gas cylinders for small-volume users. These cylinders then are loaded onto flatbed trucks for delivery. For larger-volume gas transportation, long-length gas cylinders are permanently mounted on trailer frames. These high-pressure-cylinder vehicles commonly are referred to as tube trailers. The cylinders are constructed of alloy steels, and are designed, built, tested, and maintained in accordance with U.S. Department of Transportation (DOT) regulations. The typical small-volume cylinders are constructed to a DOT 3A or 3AA specification, and typically have rated service pressures from 2015 psig to 2640 psig and water volumes from 0.39 to 1.75 ft^3. The larger cylinders often have similar pressure ratings but water volumes of approximately 75 ft^3.

The containers used for transporting and storing the majority of cryogenic liquids are insulated cryogenic tankers, liquid cylinders, and stationary storage tanks. These containers are similar in design to a thermos bottle. There are an inner vessel, which contains the cryogenic product, and an outer vessel. The space between the two vessels is under a vacuum and usually contains a wrapped insulation material such as aluminized Mylar, or it may contain a powdered insulation such as expanded perlite. The liquid cylinders, usually 160 or 250 liters in capacity, are constructed to a DOT 4L specification and operate with pressures up to 235 psig. The cryogenic tank trucks may be constructed to a DOT MC-338 specification, and range in capacity from 3000 to 14,000 gallons, usually depending on the product. The operating pressures may range from 25 to 150 psig, and the vehicles often are equipped with cryogenic pumps used to deliver the product to higher-pressure storage tanks. The stationary storage tanks are designed and built to ASME standards, typically are constructed for pressures of 250 psig or less, and may range in capacity from 500 to 70,000 gallons.[2]

Fig. 13.2. Process diagram for nitrogen PSA (pressure swing absorption) system. (*Courtesy Air Products and Chemicals, Inc.*)

Fig. 13.3. An on-site nitrogen PSA supply system. (*Courtesy Air Products and Chemicals, Inc.*)

NITROGEN

Nitrogen makes up 78 percent of the earth's atmosphere. It is a colorless, tasteless, nontoxic, and inert gas. Nitrogen does not support combustion nor does it support respiration; but it reacts with the more active metals such as lithium and magnesium to form nitrides, and at high temperatures it will combine with hydrogen, oxygen, and other elements. As a cryogenic liquid, nitrogen is nonmagnetic, stable against mechanical shock, and free of toxic or irritant vapor.

Nitrogen usually is produced by cryogenic air separation, liquefaction of atmospheric air, and removal of the nitrogen by fractionation, as described earlier. Noncryogenic adsorption technology, including PSA and VSA systems, economically provides nitrogen for requirements between 4000 and 30,000 standard cubic feet per hour (scfh).

Nitrogen has many applications, both as a gas and as a liquid. As a gas, nitrogen is employed primarily as a fluid for inerting and pressurizing systems, used, for example, in the following ways: for the agitation of color film solutions in photographic processing and for blanketing of oxygen-sensitive liquids and volatile liquid chemicals; in the production of semiconductor electronic components and for dressing nonferrous metals; in the transportation industry, in pressurizing aircraft tires and emergency bottles to operate landing gear; and often to purge and pressurize pipelines, reactor vessels, or storage tanks throughout the petrochemical industry.

Gaseous and liquid nitrogen has applications in the foods industry in the areas of food processing and packaging. Liquid nitrogen is used for rapid freezing of expensive or highly perishable foods, such as shrimp and other meat products, as well as to refrigerate foods in long-distance hauling. A gaseous nitrogen atmosphere can help to retard the spoilage of some food products.

Liquid nitrogen is used in the cold-trapping of material such as carbon dioxide from gas streams, as a coolant for electronic equipment, for pulverizing plastics or rubber materials, and for simulating the conditions of outer space. It is used to freeze liquids in pipelines during repairs to the pipeline. Liquid nitrogen frequently is employed in medicine to preserve whole blood, livestock sperm, and other biological specimens.

Liquid nitrogen gas is transported in liquid cylinders, nonpressurized dewars, cryogenic tankers, and trailers. Gaseous nitrogen is transported in high-pressure gas cylinders and tube trailers. Both forms of nitrogen also are transported via pipeline when locations relative to sources permit.

OXYGEN

This substance, which occupies 21 percent of the earth's atmosphere, has a number of very important uses. Oxygen is a colorless, odorless, and tasteless gas that is essential to the support of life. All elements except the inert gases combine directly with oxygen to form oxides. Oxygen is nonflammable, but is an oxidizer that readily supports combustion.

A large percentage of the oxygen produced commercially is made by cryogenic air separation. However, this separation is not the only oxygen-producing system; oxygen also may be made by the most common decomposition reaction, electrolysis of water. This reaction produces not only oxygen but hydrogen as well, although it is not usually an economically viable alternative when compared to other technologies.

Oxygen's major applications stem from its life-supporting and combustion-supporting properties. It is used extensively in medical applications for therapeutic purposes, for resuscitation in asphyxia, and with other gases in anesthesia. Also it is used in high-altitude flying and deep sea diving, and it is an inhalant and a fuel oxidizer in the U.S. space program.

For industrial applications, oxygen is in large demand for use with acetylene, hydrogen, and other fuel gases for purposes such as metal cutting, welding, hardening, and dehydrating. Oxygen enhances the capability of steel and iron furnaces. Oxygen is used in the production of synthesis gas—hydrogen–carbon dioxide mixtures from coal, natural gas, or liquid fuels; the synthesis gases are used to make gasoline,

Fig. 13.4. Oxy-fuel burner firing into an arc furnace. (*Courtesy Air Products and Chemicals, Inc.*)

methanol, and ammonia. Oxygen may be used for the cracking of methane or natural gas by partial oxidation to produce acetylene, an important basic chemical.

One of the largest commercial applications for oxygen is the Basic Oxygen Furnace (BOF) for refining iron and making steel. Another common application is its use in oxy-fuel burners in the steel industry. It is used in electric arc furnaces, which melt scrap metals by passing an electric current through large graphite electrodes. Figure 13.4 shows such a burner.

Oxygen enrichment is a process whereby oxygen is added to an air stream to improve a plant's efficiency. Combustion results in the generation of heat and gaseous by-products, usually carbon monoxide, carbon dioxide, and water vapor. Oxygen enrichment is used in pulp mills to increase lime production from existing equipment and eliminate the need to purchase lime from other industries. It is used again in the pulp mills to reduce airborne sulfur emission to comply with federal environment regulations.

Gaseous oxygen is transported in high pressure gas cylinders, tube trailers, and portable cylinder modules. Liquid oxygen is transported as a cryogenic fluid in liquid cylinders and insulated cyogenic tankers. Customer on-site storage is usually accomplished in large insulated cryogenic tanks.

ARGON

A relatively scarce gas, argon is colorless, odorless, tasteless, and nontoxic. It forms no known chemical compound; thus it is extremely inert.

Representing only 0.93 percent of the earth's atmosphere, argon is utilized in a variety of applications. In the manufacture of lighting devices, it is used to fill lamp bulbs and to produce colors in display tubes for lighted signs. In the field of electronics, argon provides an inert atmosphere for the production of semiconductors. Some metal welding and cutting applications are required to have an inert gas shield, and argon gas often is used

AIR SEPARATION PLANT

Fig. 13.5. Flow diagram for an air separation plant. (*Courtesy Air Products and Chemicals, Inc.*)

to fill this requirement. Also it is used as a blanket in the production of many reactive materials, including titanium and zirconium. It is often used to provide a protective shield for the growing of silicon and germanium crystals.

Argon is manufactured in oxygen–nitrogen plants by means of fractional distillation after the liquefaction of air as described above, in the cryogenic air separation process. In the distillation process where air is separated into oxygen and nitrogen, a stream of gas is withdrawn from the low-pressure column to enter the crude argon column. The argon content of the gas, initially 10 percent, is increased to 95 percent by subjecting the gas to several stages of distillation. This "crude" argon is further purified to yield the final pure argon product. Figure 13.5 shows a typical air separation process flow diagram.

For economic reasons argon often is shipped as a liquid. Liquid argon commonly is transported in liquid cylinders, cryogenic trailers, and portable tanks. Argon can be shipped in high-pressure gas cylinders and high-pressure cylinder trailers.

HYDROGEN

Hydrogen, the lightest element known, is extremely flammable and is found in 0.00005 percent concentration in air. The hydrogen molecule exists in two distinct forms: ortho and para, named according to their types of nuclear spins. Ortho-hydrogen molecules have a parallel spin, and para-hydrogen molecules have an anti-parallel spin. There is no difference in the chemical or transport properties of these forms, but there is a difference in thermodynamic properties. Para-hydrogen is the form preferred for rocket fuels. Hydrogen consists of about three parts ortho and one part para as a gas at room temperature, which is referred to as normal or equilibrium hydrogen. The concentration of the para form increases with decreasing temperature until, for the liquid, the para concentration is nearly 100 percent. If hydrogen should be cooled and

liquefied rapidly, the relative three-to-one concentration of ortho to para would not immediately change.[2]

Conversion to the para form takes place at a relatively slow rate and is accompanied by the release of heat. For each pound of rapidly cooled liquid hydrogen that changes to the para form, enough heat is liberated to vaporize approximately 1.5 lb of liquid hydrogen. However, if a catalyst is used in the liquefaction cycle, para-hydrogen can be produced directly without loss from self-generated heat.

Hydrogen can be produced by steam methane reformation (the most widely used process) or by off-gas purification. Hydrogen can be produced at normal purity (95–98%) or at ultra-high purity (99.99%). The first step to obtain purity hydrogen in the steam reformer process when using a natural gas feedstock is to perform a pretreatment. There are small amounts of sulfur in natural gas that must be removed. The sulfur removal is done by preheating the natural gas to 700°F, and then sending it through a catalytic reactor. The catalyst inside the reactor absorbs the sulfur. The second step is steam methane reforming. In this procedure, natural gas and steam are sent through a set of reformer tubes which are packed with a reforming catalyst. The reaction is heated by a furnace inside the reformer. The methane and steam react, and the results are 45 percent hydrogen, 40 percent steam, 8 percent carbon monoxide, and 7 percent carbon dioxide. The third step to obtain normal purity hydrogen is gas shifting. Carbon monoxide, hydrogen, and steam are sent through a shift converter where the gas is heated between 400°F and 700°F. The carbon monoxide and steam react to yield hydrogen and carbon dioxide.

To produce pure hydrogen, the carbon dioxide must be removed. The gas passes through a carbon dioxide removal system, which contains a chemical solvent that goes through the gas and selectively absorbs the carbon dioxide. Heat then is added to the solvent to discharge the carbon dioxide. The regenerated solvent is returned to the system to continue the removal of carbon dioxide.

To obtain ultra-pure hydrogen, the pressure swing adsorption (PSA) system is considered. In this process, hydrogen, steam, and carbon dioxide are introduced to the system. At high pressure, the gas is passed through an adsorbent bed, and the adsorbent picks up the impurities and allows hydrogen to pass through the bed. In the regeneration process, the pressure is lowered to purge out impurities. Then the hydrogen is used to rinse the adsorbent and repressurize the vessel. The efficiency of the hydrogen purification system is optimized at an operating pressure that allows use of a catalyst tube in the reformer.

There are four other distinct processes by which hydrogen may be produced under the off-gas purification method: pressure swing adsorption (PSA), membrane separation, catalytic purification, and cryogenic separation. The cryogenic separation process uses cryogenic principles to condense and separate the impurities from hydrogen. Impure hydrogen is introduced to the cryogenic system or "cold box." The impurities are factored out, and pure hydrogen is obtained. Figure 13.6 shows a typical facility where liquid hydrogen is stored and produced.

Membrane separation makes use of two layers of membrane material; the first layer

Fig. 13.6. Liquid hydrogen plant. (*Courtesy Air Products and Chemicals, Inc.*)

is nonporous, and the second layer is porous. A gas mixture under increased pressure is subjected to the surface, which allows a smaller molecule such as hydrogen to permeate faster than a larger molecule such as carbon dioxide.

The catalytic purification process requires the use of many catalysts to convert impurities into compounds. Hydrogen, chlorides, oxygen, and carbon oxides are passed through several catalytic reactors. In the first reactor, the chlorides are absorbed. In the second reactor, oxygen and hydrogen react to obtain water/steam, which is easily separated. In the third reactor, the carbon oxides react with hydrogen to yield methane, which would be considered an acceptable impurity in this particular hydrogen product stream.

Some other sources for hydrogen are the dissociation of ammonia, steam reforming of light hydrocarbons, and by-product streams of chemical and petroleum processing. The "steam iron" process is another method used to make hydrogen. Passing steam over heated, spongy iron reduces the steam to hydrogen with additional formation of iron oxide. There are several varieties of this process. The water–gas reaction is a reaction of steam with incandescent coke or coal; it is a source of hydrogen with carbon monoxide as an additional product. In the catalytic version of this reaction, excess steam breaks down to form more hydrogen while oxidizing the carbon monoxide to form carbon dioxide. Hydrogen frequently is obtained as a by-product of cracking operations using petroleum liquids or vapors as feedstock. Also hydrogen may be produced through the electrolysis of water.

There are many applications for hydrogen. It is used to make fertilizers, as the active ingredient in detergents, and in the manufacture of polyurethanes. It finds application in the production of semiconductor devices, in the process of refining crude oil, in food processing for edible oils and shortening, and as fuel for the space shuttle. The hydrogenation of edible oils in soybeans, fish, cottonseed, and corn produces solids used in shortening and other foods. Many alcohols also are produced by the hydrogenation of the corresponding acids and aldehydes. It is possible to use hydrogen in low concentrations as a physiologically inert gas for a breathing atmosphere where multiple gases are involved.

In the metals industry, hydrogen is used for applications such as annealing metals. It is also used as a fuel in underwater oxy-hydrogen torches and with oxy-hydrogen welding and cutting systems. Hydrogen serves as a nonoxidizing shield alone or with other gases in furnace brazing and in welding aluminum, magnesium, and lead. The temperature in the oxy-hydrogen flames is about 4000°F, which is suitable for low-temperature welding and brazing. It also is used in underwater cutting. Oxy-hydrogen flames are used in the fabrication of quartz and glass. Atomic hydrogen welding is suitable for very thin stock and can be used with practically all nonferrous alloys. In the process, an arc with a temperature of about 11,000°F is maintained between two nonconsumable metal electrodes. Molecular hydrogen fed into the arc is transformed into atomic hydrogen, which sends heat from the arc to the weld zone. At the surface of the welding area, the atomic hydrogen recombines to molecular hydrogen, with the release of heat.

Liquid hydrogen is very important as a fuel for powering missiles and rockets. It is used in laboratory research on the properties of some materials at cryogenic temperatures, work often associated with studies of superconductivity. Liquid hydrogen can be used as a fuel, for either propulsion or heating.

Hydrogen gas is shipped in high-pressure gas cylinders and high-pressure cylinder trailers. Liquid hydrogen is shipped in insulated portable containers, and in cryogenic tankers. The normal liquid hydrogen supply system utilizes a cryogenic stationary storage tank rated at 150 psig, just below the critical pressure of hydrogen. A cryogenic pumping station, which has a liquid hydrogen storage tank, a high-pressure cryogenic pump capable of supplying liquid hydrogen at 3000 psig, and high-pressure gas storage tubes, often is used to supply large quantities at pressures above 150 psig.

HELIUM

Helium was discovered during an eclipse as an orange line in the spectrum of the sun's atmosphere. Physically, helium is colorless, odorless, and tasteless; and it is the second lightest element, second only to hydrogen. Helium is chemically inert and has only a 5 ppm concentration in the atmosphere.

One of helium's first uses was as a nonflammable replacement for hydrogen for inflation of lighter-than-air aircraft. Today liquid helium is vitally important in cryogenic research, as it is the only known substance to remain fluid at temperatures near absolute zero. It has a unique use as a refrigerant in cryogenics and is the only nuclear reactor coolant that does not become radioactive. Liquid helium is used extensively with superconducting magnets; in the medical field, it is used to cool the superconducting magnets for magnetic resonance imaging (MRI).

Helium has a variety of other uses in welding and lighting. It is used as an inert gas shield in arc welding, for filling cold weather fluorescent lamps, and to trace leaks in refrigeration and other closed systems. Also helium is used for specialized purging and pressurizing applications. Helium can be used in place of or together with nitrogen as the inert portion of the atmosphere in sealed environments such as space vehicles, deep-sea submergence vehicles, and diving suits. The use of helium in a plasma arc furnace, is becoming an increasingly popular method for melting precious metal ores and scrap.

Small amounts of helium are extracted from the atmosphere by fractionation methods, but not commercially because of the small amount of helium in the atmosphere. A number of natural gas wells contain helium, which can be recovered by a liquefaction and stripping process. Natural gas containing at least 0.2 percent helium has been found in the American Southwest, where the natural gas fields are the major U.S. source of helium. Those helium-rich fields are within 250 miles of Amarillo, Texas; other helium-bearing fields have been found in Saskatchewan, Canada, and in areas near the Black Sea. Helium cannot be synthesized; so conservation and cleanup recycle systems for "spent" gas are important means of preserving the earth's helium resources.

Gaseous helium is stored and transported in high-pressure gas cylinders, high-pressure gas trailers, and portable cylinder modules. Liquid helium is transported in liquid cylinders, portable tanks, and cryogenic tankers. Because of the extremely low temperature of liquid helium and the cost to produce it, special design considerations must be given to storage containers in order to maintain the helium in the liquid phase. Some containers are designed to include a sacrificial liquid nitrogen shield that intercepts heat before it is transferred into the liquid helium.

CARBON DIOXIDE—CO_2

The carbon dioxide market in the United States is served by both gaseous and liquid/solid carbon dioxide. The consumption of CO_2 gas is approximately 22 million tons per year with the enhanced oil recovery and urea production markets being the primary consumers. Liquid and solid consumption is about 5 million tons per year serving primarily the food and beverage processing industries. Carbon dioxide is used extensively in a variety of areas. It is used for pressurizing, and as a source of bubbles and the acidic taste in soft drinks. Carbon dioxide is used to fill a type of fire extinguisher that literally depends on the inertness of the compound. Carbon dioxide is also used in freezing specialty and quality foods.

More specifically, each physical phase of carbon dioxide has uses in many areas. Gaseous carbon dioxide is used for the pH control in water treatment and as a growth stimulant for plant life. Liquid carbon dioxide is used in one case as an expendable refrigerant for low-temperature testing of aviation, missiles, and electronic components. It is also used in controlling chemical reactions and for stimulation of oil and gas wells. Solid carbon dioxide, commonly called "dry ice," is used extensively as a refrigerant

for dairy products, meat products, and other frozen foods while in transit. It is also used as a cooling agent in many industrial processes such as grinding heat-sensitive materials, cold-treating metals, shrink fitting machinery parts, in vacuum cold traps, in cryo-surgery, and specimen preservation and storage.

Carbon dioxide is usually nonreactive and nontoxic. At normal atmospheric pressure and temperature, it is colorless and odorless. Carbon dioxide will not burn, nor will it support combustion. Unrefined carbon dioxide gas is typically obtained from the combustion reaction of coal, coke, natural gas, and other carboniferous fuels. The gas obtained from the combustion process is liquefied and purified by several different processes to a purity of about 99.9 percent.

The major source of carbon dioxide is as a by-product of steam-methane reforming. The resulting reformer syngas may utilize one of the following three processes in the production of carbon dioxide. Carbon dioxide may be manufactured at 95 percent purity by the membrane separation process from a similar gas or from natural gas. In the recovery of high purity carbon dioxide from the syngas or natural gas, and before any separations can occur, a pretreatment step removes any contaminants and any condensate liquids that may be present. The feed gas then enters the membrane system at 300 to 1000 psig. The permeate gas is delivered at pressures between 50 to 150 psig and at a carbon dioxide purity of about 95 percent. The high pressure residual gas containing 10 to 30 percent carbon dioxide enters a conventional treatment process such as monoethanolamine or diethanolamine for a final cleanup. The carbon dioxide off-gas from the conventional treatment process is then elevated to the same pressure as that of the permeate gas from the membrane system. The combined streams may be further compressed to higher pressures, such as 2000 psig for applications in carbon dioxide flooding in enhanced oil recovery.

Fig. 13.7. Process diagram for CO_2 recovery using solvent absorption method. (*Courtesy Air Products and Chemicals, Inc.*)

A second process used in the production of carbon dioxide is pressure swing adsorption. The feed gas usually contains approximately 20 percent carbon dioxide, 70 percent hydrogen, and the remainder methane, carbon monoxide, nitrogen, and water. The feed gas is typically under a pressure of 125 psig to 400 psig at temperatures of 80 to 120 degrees Fahrenheit. The carbon dioxide and water are strongly adsorbed in the absorb beds and the residual gas stream is depressurized for further recovery. The absorber vessel is then evacuated through vacuum blowers where the carbon dioxide, which has been adsorbed by the bed, is relased at purities essentially 99+% pure.

Recovery of pure carbon dioxide from reformer syngas has often been accomplished through absorption by liquid solvents. The solvents used are: monoethanolamine (MEA), diethanolamine (DEA), and methyldiethanolamine (MDEA). See Fig. 13.7 for the process flow diagram. In this process, the syngas stream from the reformer is sent to the stripper column where the carbon dioxide is absorbed by the solvent. The residual gas stream components are discharged from the stripper column. The carbon dioxide in solvent is sent to the regeneration column where the carbon dioxide is regenerated and discharged for further purification by compression dehydration and refrigeration steps. The solvent is returned to the stripper column. The final product is essentially 99.9 percent pure.

Liquid carbon dioxide is transported in high pressure gas cylinders or liquid cylinders, and in foam insulated tank trucks. Solid carbon dioxide is packaged in various blocks wrapped in heavy paper or heavy insulated bags. Liquid carbon dioxide is normally stored in a foam insulated tank that has a refrigeration unit to maintain the tanks temperature and pressure.

LIQUEFIED NATURAL GAS

The advantages of liquefied natural gas (LNG) are evident in that over 600 standard cubic feet of natural gas occupy only one cubic foot of space in its liquid form at −260°F and atmospheric pressure. This significant volume reduction reduces the storage and transport volume, and allows shipment of natural gas in areas where pipelines were never feasible. The primary vehicle for transporting LNG is the cryogenic tanker. It also is transported in specially designed ships.

LNG is composed predominantly of methane, which may contain minor quantities of ethane, propane, nitrogen, helium, and other components normally found in natural gas.

In the past the liquefaction of natural gas used a classic cascade cycle. The process required 120,000 horsepower for liquefaction of over 150 million standard cubic feet per day. Provisions are made for some of these cycles to use seawater for cooling. Later, baseload LNG plants utilized mixed refrigerant cycles, such as Air Products and Chemicals, Inc.'s propane precooled mixed refrigerant system. Baseload plant capacities range from about 70 mmscf/day to about 350 mmscf/day of LNG. Baseload plants move LNG from remote sites by ship to populated areas. For example, Indonesia supplies LNG to Japan for electric power generation.[7]

At baseload LNG production plants, natural gas enters the plant at near 85°F and 600 psig. The gas is treated to remove carbon dioxide and any sulfur compounds, and it is cooled with high-level propane refrigerant to remove as much water as possible. Dehydration to a dew point of about −100°F is done by carrying out regeneration with dried outlet gas by way of a high-pressure steam heater and a regeneration gas compressor. The dry, treated gas is passed through a stationary bed of activated carbon for removal of mercury, and is cooled with two successively lower levels of propane refrigeration. A scrub column makes a rough fractionation to remove the heavier and medium carbon compounds. The gas at the bottom of the scrub column flows to a fractionation train, which consists of a deethanizer, a depropanizer, and a deubtanizer where relatively pure ethane and propane are produced. Excess ethane, propane, and butane from the

Fig. 13.8. Schematic of process flow for a baseload LNG plant. (*Courtesy Air Products and Chemicals, Inc.*)

fractionation process are reinjected into the main gas stream just prior to liquefaction. The overhead from the scrub column flows to the main heat exchanger for liquefaction to LNG. Figure 13.8 shows the process flow for a LNG plant.

For years LNG has been used for peak-shaving purposes, which is the storage of excess capacity of liquid natural gas to be revaporized for the coldest days of the year. In recent years LNG has been used in heating and vehicle propulsion. Also LNG is used to produce carbon black, which has applications in the manufacture of rubber products and printing ink. The burning of high-purity LNG (methane) is done to make carbon black for particular use in electronic devices.

ACETYLENE

Acetylene (C_2H_2) is a colorless and flammable gas. At 100 percent purity, acetylene is odorless, but at commercial purity it has a garliclike odor. Acetylene can be liquefied and solidified with ease, although in both cases it explodes with extreme violence when ignited.

Acetylene can be inhaled in rather high concentrations without chronic, harmful effects. In fact, it has been used as an anesthetic. However, it is a simple asphyxiant if present in concentrations high enough to deprive the lungs of oxygen and produce suffocation.[1]

A primary method for the manufacture of acetylene is to react calcium carbide, the principal raw material, with water to produce acetylene with calcium hydroxide as a by-product:

$$CaC_2 + H_2O \rightarrow C_2H_2 + Ca(OH)_2$$

Acetylene also is manufactured by thermal or arc cracking of hydrocarbons and by a process utilizing the partial combustion of methane with oxygen. The acetylene produced from calcium carbide is basically pure, with only water and air as impurities.

Eighty percent of the acetylene produced is used for chemical synthesis. Acetylene is an important raw material for a series of organic compounds such as acetaldehyde, acetic acid, acetic anhydride, acetone, and vinyl chloride. Those compounds are used to manufacture a large group of products, including plastics, synthetic rubber, dyestuffs, solvents, and pharmaceuticals, and in the manufacture of carbon black. The other 20 percent is used for oxy-acetylene welding, cutting, heat treating, lighting purposes, buoys and beacons.

In the primary method for producing acetylene, calcium carbide is fed into a generator containing water. The resulting wet acetylene gas at 120°F to 140°F is piped to a cooler condenser where it is cooled to 80°F to 90°F and the water vapor is condensed. This is the first step in the drying process. The gas then goes to a low-pressure dryer filled with calcium chloride. The calcium chloride absorbs the moisture from the acetylene gas, which then enters a compressor, where the pressure is increased to 200 to 375 psig. An oil separator filters out excess impurities and oil. The gas next goes to a high-pressure dryer, where all remaining water is removed. The gas is then pumped into cylinders for shipment. The by-product calcium hydroxide is sent to a storage tank, where water is decanted off the top, and is recovered.

The only method of distributing acetylene other than pipeline, is by means of portable steel tanks containing a porous solid filler saturated with acetone or other suitable solvent, in which the acetylene is dissolved under pressure. Acetylene alone is not handled at pressures higher than two atmospheres because of its tendency to decompose explosively dissolved in acetone it may be stored at pressures of 10 to 15 atmospheres.

NITROUS OXIDE

Nitrous oxide (N_2O), commonly known as laughing gas, is a colorless and practically odorless, tasteless, nontoxic gas. Nitrous oxide is an oxidizer that will support combustion, but it is a nonflammable gas. It is water-soluble; but when it is dissolved, its acidity does not change.

The primary commercial method for producing N_2O is by thermal decomposition of ammonium nitrate, with nitrous oxide and water in the primary reaction:

$$NH_4NO_3 \rightarrow N_2O + 2H_2O$$

The impurities formed are mostly high oxides of nitrogen and are highly toxic. After the water is removed in a condenser, the gas is washed in a solution of potassium dichromate to remove nitric oxide, in caustic to remove nitric acid, and finally in water.

Nitrous oxide is an inhalant type of anesthetic or analgesic gas. It serves as a propellant for various aerosol products, particularly with foods such as whipped cream; and it is used as an oxidizing agent in blowtorches and as a chemical reagent in the manufacture of various compounds, both organic and inorganic. It also serves as an ingredient in rocket fuel and high performance vehicle formulations, and it is used as part of the working fluid in hypersonic wind tunnels that have been investigated recently.

Nitrous oxide is filled and transported in its liquid phase in high-pressure gas cylinders or in liquid cylinders. It is transported as a liquefied compressed gas under high pressure in cylinders and at lower pressures and reduced temperatures in refrigerated cargo tanks and insulated portable tanks. Nitrous oxide is stored in a foam-insulated tank accompanied by a refrigeration unit, which is similar to the tank used for carbon dioxide.

REFERENCES

1. Compressed Gas Association, Inc., *Handbook of Compressed Gases*, Van Nostrand Reinhold, New York, 1990.
2. Compressed Gas Association, Inc., *Handbook of Compressed Gases*, Van Nostrand Reinhold, New York, 1966.
3. *Air Products Industrial Gases Fact Book and Safe Practices Guide*, Air Products and Chemicals, Inc., Allentown, PA, 1985.
4. Vance, R., and Weinstock, H., *Applications of Cryogenic Technology*, Tinnon-Brown, Inc., Los Angeles, CA, 1969.
5. Technical Orientation Program, Air Products and Chemicals, Inc., Allentown, PA, 1986.
6. Schnell, W., Houston, C. D., and Hopper, W., "Membranes Can Efficiently Separate CO_2 From Mixtures," *Oil & Gas Journal*, **81** (33), 52–56 (Aug. 15, 1983).
7. Speir, G. A., "Indonesia's Badak LNG Project Sets New Records," *Pipeline & Gas Journal* (205), 22–24, 26, 28 (June 1978).

14

Sulfur and Sulfuric Acid

Robin W. Strickland, Ph.D.*

SULFUR

Sulfur is one of the few elements that is found in its elemental form in nature. Typical sulfur deposits occur in sedimentary limestone/gypsum formations, in limestone/anhydrite formations associated with salt domes, or in volcanic rock.[4] A yellow solid at normal temperatures, sulfur becomes progressively lighter in color at lower temperatures and is almost white at the temperature of liquid air. It melts at 114 to 119°C (depending on crystalline form) to a transparent light yellow liquid, becoming deep orange as the temperature is increased. The low viscosity of the liquid begins to rise sharply above 160°C, peaking at 93 Pa · s at 188°C, and then falling as the temperature continues to rise to its boiling point of 445°C. This and other anomalous properties of the liquid state are due to equilibria between the various molecular species of sulfur, which include small rings and chains.

Sulfur also is found as sulfide minerals in combination with iron or base metals (e.g., pyrites) and as sulfates in combination with alkali metals and alkaline earths (e.g., gypsum). Hydrogen sulfide with its "rotten egg" odor is the primary sour component of sour gas. Crude oil and coal contain a variety of complex sulfur-containing organic species. These sulfur compounds are removed from the liquid fuels by treatment with hydrogen to convert the sulfur to hydrogen sulfide, which is taken off in the gas stream. The recovery of sulfur values from sour fuels for environmental reasons is the largest source of sulfur today.

World sulfur production in 1989 was 40 million tons.[24] Over 99 percent of the sulfur that is marketed is sold as crude sulfur. The two primary grades are "bright," which is bright yellow and at least 99.5 percent (typically 99.9 + %) pure, and "dark," which can contain upwards of 0.25 percent carbonaceous material. Recovered sulfur generally is bright, whereas mined sulfur often is dark owing to traces of crude oil, which is associated with salt-dome sulfur deposits. Volcanic sulfur deposits yield bright sulfur, but it often is contaminated with toxic metal oxides. Sulfuric acid production accounts for 93 percent of the elemental sulfur

*Freeport Research and Engineering (Affiliate of Freeport-McMoran).

used in the United States. Small quantities of several specialty sulfurs are produced for a wide variety of other applications, including fumigation, bleaching, pharmaceuticals, pyrotechnics, rubber making, and cutting oils.

For many years, the standard industry practice was to ship sulfur as a crushed bulk solid. The shipment of dry bulk sulfur, however, is prone to introduce contamination and to generate sulfur dust during handling.[9] The dust presents a potential environmental hazard and can result in dust explosions if proper preventive measures are not taken. In order to minimize these problems, the modern trend in handling solid sulfur is first to form it into slate, prills, pellets, or granules.[12,28] Canada is a major shipper of formed solid sulfur, and such sulfur is used by many industries in Europe and Asia. In the United States, over 95 percent of crude sulfur is shipped molten in specially designed railcars, tank barges, tank trucks, and ships. Liquid sulfur shipment and storage also is practiced in Europe. Molten sulfur generally is preferred by the sulfuric acid industry, primarily because there is no need for remelting prior to use.

Development of the Sulfur Industry

Early humans doubtless found elemental sulfur in volcanic craters, encrusting the edges of hot sulfur springs, and embedded in limestone formations. They discovered that it would burn and used it for medicinal purposes, as a bleach, as a fumigant, as a colorant, and as incense. Its use for these purposes is mentioned in ancient writings. The Romans produced incendiary weapons from sulfur. In the thirteenth century, the Chinese invented gunpowder using sulfur, nitrate, and charcoal.

The earliest commercial sulfur came from limestone deposits, of which those in Sicily and the Italian mainland developed world markets in the eighteenth and nineteenth centuries. Traditional mining methods were used to produce sulfur ore, which was burned slowly in a pile (Calcarone) to yield crude sulfur.[40] Steam smelting in autoclaves came into use about 1859. In 1890 Gill built a multichamber furnace to improve the process's production rate and efficiency. Italian monopoly of the sulfur markets continued until the early 1900s when the Frasch process brought previously unrecoverable sulfur deposits on the North American Gulf Coast into production.

Oil exploration efforts in Texas and Louisiana in the late 1800s uncovered sulfur deposits in limestone at depths of 200 to 300 meters. Mining was complicated by intervening layers of quicksand and the presence of hydrogen sulfide gas. Numerous conventional mining attempts at Sulphur, LA, proved disastrous. Finally, in December 1894, Hermann Frasch demonstrated the hot water process for mining underground sulfur deposits. With its favorable economics, the Frasch process completely displaced the Italian sulfur industry. Ready availability of low-cost sulfur opened the way for commercial sulfuric acid production by burning sulfur. This process largely supplanted the long-standing iron pyrite combustion process for sulfuric acid production by eliminating its extensive gas-cleaning operations.

In 1883 C. F. Claus patented a process for producing sulfur from hydrogen sulfide through partial combustion over an iron oxide catalyst. A number of improvements on the Claus process were made over the next 50 years, but it enjoyed limited success as a method for producing sulfur. Its primary driving force was, and remains, the need for a means to remove the sour component of sour gas for processing reasons and for environmental compatibility. Since the 1950s the demand for larger quantities of cleaner natural gas to meet increasingly more stringent environmental standards has caused a proliferation of the family of Claus processes. The processing of sour gases has increased to the extent that the production of recovered sulfur now exceeds that of mined sulfur.

Sulfur has been produced from sulfide ores (pyrites) by thermal decomposition in the absence of air, by roasting/smelting under reducing conditions, or by reaction of the ore with SO_2. Hydrometallurgical processes have produced sulfur from metal pyrites as a by-product. These processes do not contribute significant quantities of sulfur to the world markets. The reduction of sulfate ores (e.g.,

gypsum) to sulfur suffers from high energy costs. In general, the preferred product for these processes is sulfuric acid.

Sulfur Production Processes

At the present time the Frasch process is the most economical method for extracting sulfur from native deposits. Certain constraints on the geological formations required for the Frasch process limit its use to deposits along the Gulf of Mexico, in Poland, in the Former Soviet Union and in Iraq. Other sulfur deposits may yield to the Frasch process, but they have marginal economics. Figure 14.1 shows the structure needed for Frasch mining from salt domes. The sulfur-bearing limestone must have sufficient porosity to allow the sulfur to migrate upon melting. Both the caprock and the underlying anhydrite formation should be impervious to prevent loss of the hot water pumped into the mine. These salt-dome sulfur deposits typically are lens-shaped and are 1 to 75 meters thick with diameters of a few hundred meters up to several kilometers.

A sulfur well consists of a casing and three concentric pipes reaching into the sulfur-bearing strata. The outer 8- to 10-inch pipe carries 165°C water pumped into the formation to melt the sulfur. An inner cement-lined 3- to 6-inch pipe is used to transport the sulfur to the surface. Compressed air is passed through the 1-inch tube in the center to air-lift the sulfur. Without the air lift the molten sulfur would rise only partway in the middle pipe. The compressed air produces a low-density sulfur froth that rises to the surface.

The superheated water melts the sulfur in the vicinity of the well, forming a molten sulfur pool at the bottom of the well. As production continues, the formation fills with water. To continue production, bleedwells are drilled at the periphery of the formation to allow for discharge of the cooled minewater. In some mine fields, sufficient minewater is lost to the geological formation to provide for continued production. To limit minewater loss, mud

Fig. 14.1. Frasch process for mining sulfur from salt-dome formations. *(Courtesy Freeport Sulphur Co.)*

sometimes is pumped into the formation to seal major crevices.

Although most Frasch mines are located inland, since 1960 sulfur has been mined offshore in Louisiana by using drilling techniques developed for offshore oil production. Using saltwater to provide heated minewater has presented the Freeport Sulphur Company with some unique challenges in corrosion and scaling. Figure 14.2 shows an artist's depiction of Freeport's new 67-million-ton Main Pass 299 Frasch mine located 10 miles offshore in 200 feet of water near the mouth of the Mississippi River. The inset in Fig. 14.2 shows Freeport's Grand Isle sulfur mine 7 miles off the coast of Louisiana.

Sulfur produced from salt-dome structures can be quite pure, but it often contains up to 1 percent of bituminous residues, which render it dark and can make it unacceptable for sulfuric acid production. Some purification is obtained by filtering the dark sulfur through diatomaceous earth. The carbonaceous material can be formed into larger, filterable particles (Carsul) by treating the sulfur with heat or sulfuric acid. Freeport's submerged combustion distillation process was used from 1966 until 1979 to purify Frasch sulfur with up to 2 percent carbonaceous material.

Recovered Sulfur

Hydrogen sulfide is recovered from natural gas or refinery gases by absorption in a solvent or by regenerative chemical absorption.[7,13] In either case a concentrated hydrogen sulfide stream is produced that is treated further by the Claus process. A typical Claus plant has a feed stream of at least 45 percent H_2S, but with modifications can handle streams containing as little as 5 percent H_2S. For gas streams with low concentrations of hydrogen sulfide, direct conversion of the hydrogen sulfide to sulfur is accomplished in the solvent system (e.g., Stretford process).

The Claus process is based on the reaction of hydrogen sulfide with sulfur dioxide according to the highly exothermic reaction:

$$2H_2S + SO_2 \rightleftarrows 3S + 2H_2O$$

In practice the sulfur dioxide is produced in situ by partial oxidation of the hydrogen sulfide with air or oxygen in a furnace. In the split-flow arrangement, one-third of the H_2S stream is burned and then recombined with the remaining two-thirds before entering the Claus reactor. In the straight-through version, the entire H_2S stream is sent through the burner, and the extent of H_2S combustion is controlled by the air feed rate.

A flow diagram for a typical Claus process is shown in Fig. 14.3. The hydrogen sulfide is burned in a fuel-fired furnace (950–1250°C) with air to produce sulfur and a gas stream containing H_2S and SO_2. Process controls maintain the H_2S:SO_2 ratio near 2 in accordance with the stoichiometry. Heat is removed from the gas stream in a waste heat boiler to control the process gas temperature. The process gas is passed through one or more catalyst beds to convert the H_2S and SO_2 to sulfur, which is removed in condensers between each bed. High temperature shifts the equilibria toward the reactants, whereas low temperature causes sulfur condensation on the catalyst bed, leading to decreased catalyst activity. In practice the gas temperature is closely controlled for each Claus reactor, with higher temperatures at the first reactor where compounds such as COS and CS_2 are converted to S. Although efficiencies of 96 to 99 percent sulfur recovery have been demonstrated in Claus plants, it is common to have efficiencies in the 80 to 90 percent range. The conversion efficiency can be increased with multiple reactors in series. Oxygen enrichment can increase the throughput but does not improve the conversion.[22,34] Figure 14.4 shows a typical sulfur recovery plant based on the Claus process.

The tail gas from the Claus reactors may be further processed to remove any remaining sulfur compounds. This may be done by low-temperature Claus-type solid bed processes (e.g., the Sulfreen process), wet-Claus absorption/oxidation processes (e.g., Clauspol 1500 process), or hydrogenation of the off-gas to form H_2S for recycle (e.g., SCOT process). Residual sulfur compounds in the processed tail gas then are incinerated to SO_2. The residual SO_2 in the oxidized tail gas may be

Fig. 14.2. Artist's rendition of Freeport Sulphur Co.'s new Main Pass 299 sulfur mine off the Louisiana coast. Inset shows Freeport's Grand isle sulfur mine. *(Courtesy Freeport Sulphur Co.)*

Fig. 14.3. Claus process flow diagram.

Fig. 14.4. Republic Refining Co. 440 metric ton/day Claus sulfur recovery unit (left) and SCOT tailgas cleanup unit (right). *(Courtesy Ortloff Engineers, Ltd.)*

scrubbed by any of several processes (e.g., Wellman-Lord process) before being vented to the environment. It is feasible to bring the SO_2 content of the final waste gas to as low as 200 to 500 ppmv by combinations of these processes. As of 1990, U.S. EPA regulations limited stack SO_2 emissions to 2000 ppmv. Individual states may apply more stringent regulations to achieve national ambient air quality standards for criteria pollutants including SO_2.

Production and Consumption of Sulfur

The 1989 world production of sulfur in all forms is given in Table 14.1 by region and for major producing countries. Recovered sulfur accounted for 63 percent of brimstone production in 1989. Most of the recovered sulfur was derived from sour gas wells in Canada, the Soviet Union, and the United States.[41] U.S. oil refineries recovered nearly 4 million tons of sulfur from the hydrotreating of sour crudes. Canadian tar-sands comprise the major source of other recovered sulfur production. Frasch and open pit mines in southeast Poland near the Russian border lead the world in native sulfur production. Most of the Russian native sulfur production comes from Frasch and open pit mines just across the Polish–Russian border. The U.S. Frasch sulfur production comes from West Texas and the Texas–Louisiana Gulf Coast. In 1989, China, the Soviet Union, and Spain accounted for 68 percent of pyritic sulfur production, virtually all of which was produced as sulfuric acid. Sulfur production listed as "Other Sulphur" in the table was primarily sulfur dioxide from the smelting of nonferrous metal sulfide ores.

The growing importance of recovered sulfur in world sulfur markets is mirrored in the U.S. market trends shown in Fig. 14.5.[33] Involuntary sulfur production from gas and oil began to climb in the early 1970s following environmental legislation to curb sulfur dioxide emissions. Large stockpiles of bright recovered sulfur began to appear in North America and especially in Western Canada. New Frasch sulfur mines in Poland, Texas, and Louisiana had begun production in the late 1960s, adding to the overall brimstone production and

TABLE 14.1 World Production of Sulfur in 1989 (Thousands of Metric Tons)

		Recovered Sulfur						
	Frasch/ Native	*Gas*	*Oil*	*Other*	*Total Brimstone*	*Pyrite Sulfur*	*Other Sulfur*	*Total*
West Europe	45	1,727	1,648	110	3,530	2,461	1,947	7,938
France	—	636	238	—	874	—	190	1 064
Germany FR	—	952	264	110	1 326	250	458	2,034
Spain	—	—	—	—	—	938	229	1,167
East Europe	7,765	2,775	488	305	11,333	2,690	2,957	16,980
Poland	4,865	—	35	—	4,900	—	180	5,080
U.S.S.R.	2,900	2,750	250	290	6,190	2,080	1,947	10,217
North America	3,888	7,693	4,068	591	16,240	179	1,919	18,338
Canada	—	5,156	160	526	5,842	179	699	6,720
U.S.	3,888	2,537	3,908	65	10,398	—	1,220	11,618
Latin America	1,528	500	339	—	2,367	72	741	3,180
Mexico	1,528	500	50	—	2,078	—	259	2,337
Asia	1,256	2,497	2,098	—	5,851	4,650	2,495	12,996
China PR	200	60	50	—	310	4,100	556	4,966
Iraq	1,050	250	50	—	1,350	—	8	1,358
Japan	—	—	1,176	—	1,176	161	1,423	2,760
Saudi Arabia	—	1,390	200	—	1,590	—	—	1,590
Oceania	—	—	70	—	70	29	216	315
Africa	—	3	40	—	43	580	405	1,080
World Total	14,482	15,195	8,751	1,006	39,434	10,661	10,680	60,715

Source: Freeport Sulphur Marketing Department estimates.

Fig. 14.5. Trends in U.S. sulfur consumption. (Morse 1990)[32]

causing a decrease in sulfur prices, as shown in Fig. 14.6. Low sulfur prices and higher energy costs in the early 1970s resulted in the closing of several high-cost Frasch mines, and the depletion of reserves caused the shutdown of other older Frasch mines. This reduction in Frasch sulfur production in the mid-1970s, coupled with increasing sulfur demand for fertilizers, led to sulfur price increases. These price increases were accelerated in 1979 after Canadian sulfur exports were interrupted by a blockage of Vancouver harbor, which resulted from a ship colliding with a bridge. Inflation and changes in the sulfur market

Fig. 14.6. U.S. sulfur price history. (Morse 1990)[32]

TABLE 14.2 Consumption of Sulfur in the United States (Thousands of Metric Tons)

Use	*1989*	*1984*	*1979*
Agricultural chemicals	277	480	272
Petroleum refining	142	278	103
Industrial organic chemicals	122	127	166
Industrial inorganic chemicals	188	285	192
Pulp and paper	8	29	124
Other	372	813	970
Nonacid consumption	1,109	2,012	1,827
Sulfuric acid			
Domestic sulfur	8,241	6,909	7,793
Imported sulfur	2,245	2,425	1,754
Acid consumption	10,486	9,334	9,547
Total U.S. consumption	11,595	11,346	11,374

Source: Mineral Industry Surveys, "Sulfur," U.S. Department of the Interior, Bureau of Mines.

prevented a return of sulfur prices to pre-1979 levels, although substantial liquidation of sulfur inventories in the early 1980s did cause a dip in prices. Sulfur stockpiles now have been largely depleted, and sulfur prices have stabilized in recent years.

More than 90 percent of the elemental sulfur consumed in the United States is converted to sulfuric acid. Table 14.2 shows the pattern of U.S. sulfur consumption during the 1980s.[33]

SULFURIC ACID

Sulfuric acid is the largest-volume industrial chemical manufactured in the world. Over 39 million tons of sulfuric acid were produced in the United States in 1990, of which 70 percent was used in fertilizer production.[33] Its use extends to nearly every major chemical sector, and its consumption is often cited as an indicator of the general state of a nation's economy.[35] This versatile acid is truly the "work horse" of the chemical industry.

Pure sulfuric acid is an oily, water-white, slightly viscous liquid with a melting point of 10.4°C and a boiling point of 279.6°C. It is infinitely miscible with water, forming sulfuric acid solutions characterized by their weight percent H_2SO_4. Oleum may be formed by dissolving SO_3 in sulfuric acid to attain fuming sulfuric acid, with concentrations nominally greater than 100 percent H_2SO_4. Historically, sulfuric acid concentrations were determined by measuring the solution density using hydrometers calibrated in degrees Baume (°Be). This practice is waning although some specifications and tables of properties still include this measurement.

Uses of Sulfuric Acid

The primary industrial uses of sulfuric acid are in phosphate fertilizer manufacture, petroleum refining, copper ore leaching, synthetic rubber and plastics, and pulp and paper mills.[25] It is used as a solvent, a dehydrating agent, a reagent in chemical reactions or processes, an acid, a catalyst, and an absorbent and in many other applications. In spite of its wide usage, sulfuric acid rarely is contained in the final product. Sulfuric acid ends up as gypsum in phosphate fertilizer manufacture, for example. In many other processes the sulfuric acid is converted to a waste product that requires disposal or reuse. Because disposal of waste sulfuric acid is becoming increasingly unacceptable environmentally, the recycle of sulfur values from waste sulfuric acid has become more widespread.

Nearly all sulfuric acid is manufactured by the contact process in which sulfur trioxide is absorbed into 93 to 98 percent H_2SO_4. The acid may be sold at various strengths, usually depending on the requirements of the consumer. It generally is marketed on a 100 percent basis, but normally is shipped as 93 percent H_2SO_4 (66°Be), as 98 percent acid, or as 20 to 22 percent fuming oleum. Table 14.3 shows common acid strengths and end uses. Concentrated acid may be stored in mild steel tanks, but dilute acid must be contained in lead-lined or plastic tanks. Bulk shipments of concentrated acid are made in steel tanks on ships, tank barges, or railcars. Reagent grade acid is commonly sold in 5-liter glass bottles.

Development of the Sulfuric Acid Industry

Sulfuric acid is formed in nature by the oxidation and chemical decomposition of naturally ocurring sulfur and sulfur-containing

TABLE 14.3 Acid Strengths and End Uses

Percent H_2SO_4	*°Be*	*Percent Oleum (% Free* SO_3*)*	*Specific Gravity*	*Uses*
35.67	30.8	—	1.27	Storage batteries, electric utilities.
62.18–69.65	50–55	—	—	Normal superphosphate and other fertilizers.
77.67	60.0	—	1.7059	Normal superphosphate and other fertilizers; isopropyl and *sec*-butyl alcohols.
80.00	61.3	—	1.7323	Copper leaching.
93.19	66.0	—	1.8354	Phosphoric acid, titanium dioxide, steel pickling, regenerating ion exchange resins.
98–99	—	—	—	Chlorine drying, alkylation, boric acid.
104.50	—	20	1.9056	Surfactants, nitrations.
106.75	—	30	1.9412	Hydrofluoric acid.
109.00	—	40	1.9737	Explosives.
111.24	—	50	1.9900	Reagent manufacture, organic.
113.50	—	60	1.9919	Sulfonations, blending with.
114.63	—	65	1.9842	Weaker acids.

Source: *Chemical Economics Handbook*, SRI International, Dec. 1990.

compounds. It is made by the action of bacteria (*Thiobacillus ferrooxidans*) on coal brasses or iron disulfide discarded on refuse dumps at coal and copper mines; it is produced in the atmosphere by the oxidation of sulfur dioxide emitted from the combustion of coal, oil, and other substances; and it also is formed by chemical decomposition resulting from geological changes.

Although there were vague references to "spirits" expelled from alum by Arabian alchemists in the tenth century and by Roman alchemists in the thirteenth century, the first distinct mention of sulfuric acid has been credited to Basil Valentine in the late 1400s.[21] He burned sulfur with saltpeter in glass retorts or bell jars with a little water, and he also calcined copperas (ferrous sulfate heptahydrate) with silica, with both processes yielding sulfuric acid, although he took them to be different substances. These processes for making sulfuric acid continued until 1746 when Dr. Roebuck constructed a lead chamber in England for sulfuric acid manufacture. This marked the beginning of the "chamber process" for sulfuric acid, which was to continue in use for the next two centuries.[19]

The first lead chamber was 1.8 meters square, and 8:1 mixtures of sulfur and saltpeter in iron carts were rolled into it and burned with intermittent admission of air.[42] As in the glass retorts, the sulfur trioxide that was formed combined with water to produce sulfuric acid, which condensed on the walls and collected in pans. Steam was introduced into the chambers in 1774, and continuous addition of air was begun in 1793. It then was recognized that the sulfurous acid from the burning of sulfur was oxidized by air and needed saltpeter only as a catalyst. In 1827 Gay-Lussac invented a tower for recovering the nitrogen oxides escaping from the chamber. The nitrogen oxides were condensed in sulfuric acid but could not be economically recovered from the acid until the invention of the Glover denitrating tower in 1859. The introduction of these two towers completed the chamber process except for various refinements to reduce costs. The process could produce acid with up to 77 percent H_2SO_4 but generally yielded strengths in a 65 to 68 percent range. Higher-strength acid was produced by boiling chamber acid to remove water. A platinum still for producing concentrated sulfuric acid from chamber acid was first built for the Harrison Works in Philadelphia in 1814.[16]

The developing markets for sulfuric acid in the late eighteenth century increased the demand for Sicilian sulfur.[23] By 1832, sulfur prices had risen to $80 a ton, and stocks rose; then in 1833 the market broke, with the sulfur

price at $15 a ton. Sicilian government attempts to stabilize sulfur at $70 a ton failed, in part because of an 1833 discovery by a French chemist that sulfur dioxide could be obtained by roasting pyrites in a furnace. Process improvements in pyrite roasting by 1870 made pyrites competitive with sulfur as a raw material. By 1880, with the single exception of the United States, the sulfuric acid industries had gone to a pyrites basis. By 1909 virtually all U.S. sulfuric acid came from pyrites and as by-product acid from zinc and copper smelters. In the late 1890s, development of the Frasch process for sulfur mining lowered the price for sulfur. This change, coupled with new developments in sulfur burners, led sulfuric acid manufacturers back to sulfur as a raw material source. During World War I gypsum was burned with coal to produce SO_2 for sulfuric acid production. This process continued in Europe until the 1950s, with the last plant (in Great Britain) shut down in 1975. Today sulfuric acid production from pyrite roasting remains important in Spain, the Former Soviet Union, China, Japan, South Africa, Turkey, and some eastern European nations.[39] South Africa still has a gypsum sulfuric acid plant in operation.

The contact process for sulfuric acid dates from 1831, when a Briton patented a method for converting sulfur dioxide to sulfur trioxide by passing the gas through a heated tube filled with finely divided platinum.[21] The sulfur trioxide was adsorbed in chamber acid to produce concentrated and fuming sulfuric acids. Commercialization of the new process was delayed by lack of markets for the concentrated acid and poor understanding of the process parameters. About 1870, demand for fuming sulfuric acid spurred German development of the contact process. In 1901, BASF reported that the governing principles for the successful manufacture of contact acid were well understood. These included cleaning of the sulfur dioxide gas stream, use of excess oxygen, and temperature control of the catalyst bed. However, the reliance of the contact process on feed acid produced by the chamber process limited its development until about 1930. It became known that the acid in the final SO_3 absorption tower had to be kept in the range 98.5 to 99.0 percent H_2SO_4 to be effective. Accurate dilution techniques to allow recycle of the absorber acid were developed in the late 1920s. The contact process then could continuously produce diluted acid, and no longer required chamber acid. The development of poison-resistant vanadium catalysts for SO_2 conversion allowed the use of smelter gases in the contact process. With its lower capital and operating costs, the contact process then supplanted the chamber process for sulfuric acid production. By 1940 the United States had equal numbers of chamber and contact processes. Today no commercial chamber plants are operating in the United States.

Since the 1940s most developments in the contact process have focused on energy recovery and pollution abatement. The production of SO_2, its conversion to SO_3, and the dilution of H_2SO_4 are exothermic processes that are exploited to reduce energy costs at sulfuric acid plants. Sulfur emissions have been sharply reduced by using two SO_3 absorbers although one absorber is the economic choice. Tail gas scrubbing processes have been developed to further reduce sulfur emissions from sulfuric acid plants. As of 1990, U.S. EPA regulations limited sulfuric acid plant stack emissions to 2 kg SO_2 and 0.075 kg acid mist per metric ton of H_2SO_4 produced.

Manufacture of Sulfuric Acid by the Contact Process

The basic steps in the contact process are: (1) production of sulfur dioxide; (2) cooling and, for smelters, cleaning of the process gas; (3) conversion of the sulfur dioxide to sulfur trioxide; (4) cooling of the sulfur trioxide gas; and (5) absorption of the sulfur trioxide in sulfuric acid.[35] Figure 14.7 is a photograph of a contact process plant. A simplified diagram for a double absorption contact sulfuric acid process is shown in Fig. 14.8. Because sulfur dioxide is produced by several processes, it is convenient to separate the discussion of sulfur dioxide production from its conversion to sulfuric acid.

Fig. 14.7. Contact process sulfuric acid plant, showing a double absorption train (bottom) and three single absorption trains. *(Courtesy Agrico Chemical Co.)*

Sulfur Dioxide Production

Sulfur is converted to sulfur dioxide by burning molten sulfur with dried air in a sulfur burner to yield a 1000 to 1100°C gas stream containing 10 to 11 percent SO_2. The burner is mounted at one end of a sulfur furnace, and the gas passed through a waste heat boiler at the other end. The gas temperature is reduced to 420 to 440°C on leaving the boiler, which produces 40-bar steam. The gas stream usually is clean and requires no further treatment before entering the converter.

In the simple pressure-nozzle burner, the liquid sulfur is atomized by pumping it at 8 to 15 bar through the nozzle. In a two-component burner, the sulfur is atomized primarily by the combustion air stream. It operates at lower pressure and has a wider sulfur throughput range, 5 to 170 tons/day, than the simple burner. Lurgi's rotary burner can burn up to 400 tons/day of sulfur. Larger sulfuric acid plants use multiple sulfur burners. A process gas with 18 percent SO_2 can be achieved in Lurgi's two-stage sulfur combustion process. In the two-stage process, excess sulfur is burned in the first stage, consuming all of the oxygen. The low residual oxygen level limits the formation of nitrogen oxides, which otherwise would be very high at the combustion temperature of 1750°C. The combustion gas then is cooled to 620 to 650°C, and additional dried air is injected to burn the residual sulfur in an afterburner. A second waste heat boiler cools the process gas to 420 to 440°C before sending it to the converter.

Fig. 14.8. Schematic flow diagram of a double absorption contact process for sulfuric acid manufacture.

Pyrites and other iron sulfides are roasted to produce an iron oxide cinder and an off-gas containing 7 to 14 percent SO_2, which is contaminated with varying amounts of arsenic, lead, zinc, and other metal oxides. The off-gas must be cleaned before it is sent to the converter. Various types of pyrite roasting equipment have been used in the past, including shaft furnaces, multiple hearth roasters, rotary kilns, and dust roasters. Fluid-bed roasters now are widely used for their superior process technology, throughput rates, and economics. The roasting process must be controlled between 850 and 940°C. At lower temperatures the reaction is incomplete, whereas at higher temperatures the iron oxides and sulfides form a eutectic melt that inhibits the reaction rate. Fluid-bed roasters surpass other types in temperature control and temperature uniformity throughout the bed. Fluid bed technology for SO_2 production was introduced in the early 1950s by Dorr Oliver in the United States and by BASF in the Federal Republic of Germany.

Copper, lead, zinc, and other sulfide ores may be processed by roasting or smelting.[38] Roasting or sintering of sulfide ore is essentially identical with pyrite roasting. Smelting generally occurs at higher temperatures. Older reverberatory furnace smelters produce off-gas with only 1 to 2 percent SO_2, too low for its economical recovery as acid. By using an oxygen-enriched air feed, the off-gas can be raised to 6 to 8 percent SO_2. Bath smelters (Mitsubishi, Noranda), where the ore concentrates are heated and reacted in the slag/matte melt, produce an off-gas with 10 to 20 percent SO_2.[38] Flash smelters (Inco, Outokumpu), which involve suspension and reaction of the concentrates in an oxidizing gas stream, operate at 1200 to 1300°C and produce a waste gas with 10 to 15 percent SO_2. Oxygen enrichment of the feed air can raise the SO_2 level in the off-gas to 30 to 80 percent. The Kivcet process smelts with pure oxygen and produces off-gas with 80 to 85 percent SO_2.[11] Figure 14.9 is a photo of a metallurgical gas sulfuric acid plant.

Off-gas from roasting and smelting operations may contain dust, SO_3, halogens, NO_x, arsenic and other toxic metal fumes, and mercury.[35,43] These components must be

Fig. 14.9. Metallurgical gas sulfuric acid plant. *(Courtesy Texasgulf Inc.)*

removed from the gas stream before it is sent to the converter. Although SO_3 is produced in the converter, its presence in the cooled gases upstream of the converter will cause excessive corrosion by forming sulfuric acid mists.[14] The cleaning plant steps are: (1) hot-gas dedusting; (2) wet scrubbing; (3) gas cooling; (4) mist removal; and, if necessary, (5) mercury removal. The hot gases generally are passed first through a waste heat boiler to reduce the temperature to 250 to 400°C. Cyclones followed by hot-gas electrostatic precipitators (50–90 kV) remove nearly all (99+%) of the dust. The gas is then contacted with weak (5–30%) H_2SO_4 in a spray tower or, preferably, a venturi scrubber, which removes metal vapors and additional solids, cools the gas to 50 to 80°C, and converts SO_3 to acid mist for later removal. The saturated gas next is cooled in a packed bed tower or shell and graphite tube heat exchanger to condense excess water. Silica packing may be used in the towers to remove fluorides as fluosilicic acid. Acid mist is removed in wet electrostatic precipitators fitted with lead tubes and star wires or, more recently, with FRP or PVC tubes specially treated to maintain a conductive liquid film surface (Lurgi)[8] and composite wires fabricated with barbs to promote corona discharge.[43] If present, mercury is removed in an additional tower by scrubbing with mercury chloride solution (Boliden),[6] hydrochloric acid solution, or 70 to 85 percent sulfuric acid.

Waste sulfuric acid sludges from petroleum refineries are disposed of by conversion to SO_2 for production of fresh sulfuric acid. The heavy organic components of the sludges can be decomposed thermally at 800 to 1300°C (Lurgi, Monsanto Enviro-Chem) or reductively at 200 to 600°C with coal in a rotary kiln (Chemico). Thermal decomposition is accomplished in a fuel-fired vertical or horizontal furnace. The acid sludge also can be injected into fluid-bed pyrite roasters as a means of disposing of the acid and reducing fossil fuel consumption. Dilute acid sludges must be concentrated to 60 to 75 percent H_2SO_4 for economical conversion. This is generally done by using waste heat from the decomposition process. The sulfur dioxide gas stream from these processes requires cleaning as described above for roasting and smelting plants.

Calcium sulfate may be decomposed to cement clinker and sulfur dioxide gas in a coke-fired rotary kiln at 900 to 1400°C (Müller-Kühne).[35] However, the unfavorable economics of this process relegate it to countries that do not have other sources of sulfur. Phosphogypsum (gypsum produced by the acidulation of phosphate rock) may be decomposed in this way as a means of recycling the sulfur values in the large waste phosphogypsum piles at fertilizer plants (OSW-Krupp). This process is hampered by the more extesive gas cleaning requirements for decomposing phosphogypsum as compared to natural gypsum. Interest in recycling phosphogypsum remains high, driven by environmental considerations. Recently a pilot plant was constructed in Louisiana to demonstrate FIPR/Davy McKee's circular grate process for recovering SO_2 from phosphogypsum. Environmental forces also are behind the recycling of ferrous sulfate from metals industry pickling liquors. This "green salt" is decomposed to sulfur dioxide and iron dioxide in pyrite roasters. Elemental sulfur, coal, or fuel oil may be used as supplementary fuels.

Refinery waste gases may be burned to eliminate hydrogen sulfide and other sulphur-containing contaminants. Streams containing small amounts of H_2S or constituents unsuitable for Claus plants may require combustion to SO_2 as the means of disposing of the toxic gas. The resulting effluent gas usually is low in SO_2 and contains water vapor and carbon dioxide. Flue gases from fossil fuel power plants also fall into this category. Recovery of the sulfur values from these dilute gases usually is driven by environmental considerations rather than economics. In the United States, power plant flue gas often is scrubbed with lime to convert the SO_2 to gypsum sludge for landfill disposal. Alternatively, the Bergbau-Forschung process recovers sulfur dioxide by dry adsorption on activated coke at 130°C. The SO_2 is released

by heating the coke to 600 to 650°C. Sulfur dioxide can be absorbed in a sodium sulfite solution (Wellman-Lord) to produce sodium bisulfite. Pure moist sulfur dioxide can be recovered by heating the sodium bisulfite.

Conversion of SO_2 to H_2SO_4

The single absorption contact process for sulfuric acid is characterized by four main process steps: gas drying, catalytic conversion of SO_2 to SO_3, absorption of SO_3, and acid cooling. The maximum SO_2 conversion for a single absorption plant is about 97.5 to 98 percent. By adding another SO_3 absorber, the SO_2 conversion can be increased to 99.5 to 99.7 percent, resulting in lower SO_2 emissions. The so-called double absorption process is now the industry standard.

The gas drying step is not used in the wet catalysis process that was developed for waste gas streams from burning hydrogen sulfide. In the wet process, liquid acid is formed by condensation of H_2SO_4 vapor rather than by absorption of SO_3. However, the SO_2-laden process gas sent to the converter must be dry to protect the downstream process equipment against corrosion. The drying generally is done in a packed tower with recirculating concentrated (93–98%) sulfuric acid kept at 50 to 60°C by indirect cooling. The tower acid stream is heated by condensation of the water and by dilution of the acid. The tower temperature is used to control the moisture level of the gas sent to the converter. Acid from the drying tower is sent to storage tanks for shipment. When sulfur is burned, the combustion air to the sulfur burner is dried because the combustion of sulfur does not produce water. Off-gases from pyrite roasters and metallurgical smelters are dried as part of the gas cleaning process. For roaster gases with low SO_2 concentrations, a predryer may be added upstream of the main dryer.

Oxidation of SO_2 to SO_3 is accomplished in multi-stage fixed-bed catalytic converters equipped with interstage boilers or heat exchangers to remove the heat of reaction. Typically, four stages are compartmented within a single vertical converter, which may be brick-lined, steel, or cast iron. Newer converters are stainless steel, and some have five stages for higher conversion. Isothermal tubular converters are no longer suited to modern high-capacity plants. The extruded cylindrical catalyst pellets are usually 4 to 9 percent V_2O_5 with alkali metal sulfate promoters on a silica carrier (diatomaceous earth, silica gel, or zeolites). The reaction temperature is generally 410 to 440°C for these catalysts although several low-strike catalysts now operate at 360 to 400°C. Higher temperatures ($\sim$600°C) reduce the SO_2 conversion and lead to structural damage of the catalyst. High pressure drops across catalyst beds from catalyst dust formed during processing require periodic catalyst removal and screening to remove dust. Ring-shaped catalysts developed by Topsoe have lower dust pressure drops and now are in wide use. Other catalyst shapes used are ribbed rings and cylinders. The usual catalyst loading per one ton/day sulfuric acid capacity is 150 to 200 liters in a double absorption plant and 200 to 260 liters in a single absorption plant. Bayer developed and operates fluid-bed converters that utilize special 0.3 to 1 mm abrasion-resistant catalysts.

Absorption of SO_3

Sulfur trioxide from the converter is absorbed in 98 percent H_2SO_4 recirculated countercurrently through a packed tower maintained at 60 to 80°C by indirect cooling. The optimum concentration of the absorber acid is near the H_2O–H_2SO_4 azeotrope, 98.3 percent H_2SO_4, where the SO_3, H_2SO_4, and H_2O vapor pressures are at their lowest values. Absorption efficiencies in excess of 99.9 percent generally are obtained. On leaving the converter, the process gas is cooled first with feed gas in a gas–gas heat exchanger and then with a water feed preheater to 180 to 220°C before it enters the absorber. An impingement separator, or Teflon or glass fiber mist eliminator, is placed in the top of the absorber to remove acid mists. If oleum is produced at the plant, it is made in a separate oleum tower upstream of the absorber. A portion of the SO_3 stream to the absorber is diverted to the

oleum tower where it is absorbed in 98 percent acid.

In double-absorption plants an additional absorber is placed between the second and third (or between the third and fourth) converter beds. By removing SO_3 from the gas stream at this intermediate point, higher SO_2 conversions are attained in the downstream converter beds, and the overall SO_2 conversion is increased. The cooled gas from the intermediate absorber is reheated by hot converter gas in gas–gas heat exchangers before returning to the converter. An oleum tower may be placed before the intermediate absorber.

Absorption of SO_3 in concentrated sulfuric acid and the formation of H_2SO_4 from SO_3 and H_2O produce heat in the absorber, as does acid dilution from makeup water addition. Process control requires that the acid be cooled before it is recirculated to the dryer or absorber towers or sent to storage. Earlier acid coolers of parallel banks of stacked, irrigated, cast iron sections have been largely replaced by stainless steel shell and tube or plate exchangers, with or without anodic protection. Hastelloy, Sandvik SX, and Saramet alloys and Teflon also are used in acid piping and coolers.[15,17,27]

Tail gas emissions are controlled by improving the SO_2 conversion efficiency and by scrubbing the tail gas. In a double-absorption process plant, a five-bed converter has 0.3 percent unconverted SO_2, as compared to 0.5 percent for a four-bed converter. A Lurgi Peracidox scrubber may be used to remove up to 90 percent of the residual SO_2 in the tail gas from a double-absorption plant. Hydrogen peroxide or electrolytically produced peroxymonosulfuric acid is used to convert the SO_2 to H_2SO_4 in the Lurgi scrubber. Tail gas from single-absorption plants may be absorbed on activated carbon (Sulfacid) or scrubbed with ammonia or sodium sulfite (Wellman-Lord).

Metallurgical acid plants differ from sulfur-burning plants in that the cleaned SO_2 process gas must be heated before it is sent to the converter. Many of these plants have weak SO_2 streams that require large gas–gas heat exchangers for temperature control. Four plants in the Former Soviet Union processing 2 to 4 percent SO_2 use an unsteady-state oxidation process in which the cold (40–70°C) SO_2 gas is reacted on hot catalyst beds without intervening heat exchangers.[1,30] As the temperature front moves through the bed to the exit side, the flow is reversed. Cycle times are 30 to 120 minutes, and single-bed conversions of 80 to 90 percent are reported as compared to 55 to 60 percent for conventional processes at higher exit temperatures.

The thermal capacity of a 1000 ton/day sulfuric acid plant is about 63 MW. This heat liberation must be controlled in a manner that maintains optimum gas temperatures in the converter system and optimum acid temperatures in the dryer and absorber circuits. Tail gas emissions also are affected by the energy balance. Figure 14.10 shows an energy flow diagram for a contemporary sulfur-burning sulfuric acid plant. About 97 percent of the total energy input derives from burning sulfur, and 3 percent comes from electricity consumed to drive the gases through the plant. Most plants can recover 55 to 60 percent of the energy as high-pressure steam (40 bar, 400°C), but about 40 percent is lost as waste heat dissipated to the environment in the form of hot water from acid coolers.

During the late 1970s acid plants were optimized to generate more steam. Steam can be produced at pressures up to 80 bar from high-temperature sulfur burners such as the Lurgi two-stage combustion system. The high pressure steam is let down to low pressure steam through a turbo generator that cogenerates electricity. In double absorption plants, economizers were installed upstream or downstream of the heat exchanger servicing the intermediate absorber. Boiler feedwater is preheated to 90 to 95°C in this economizer to increase steam production. The energy recovery from acid plants was increased to 70 percent by installing low gas-temperature economizers, low pressure drop catalysts, and suction drying towers, by increasing the SO_2 feed gas concentration, and by preheating the boiler feedwater with hot acid. Further energy

A BLOWER; B SULPHUR FURNACE; C WASTE HEAT BOILER; D CATALYST BED 1; E STEAM SUPERHEATER; F CATALYST BED 2; G BOILER; H CATALYST BED 3; J INTERMEDIATE HEAT EXCHANGERS; K INTERMEDIATE ABSORBER; L CONVERTER BED 4; M ECONOMIZER; N FINAL ABSORBER; O AIR DRYER; P ACID COOLERS

Fig. 14.10. Sankey energy flow diagram for a 1000 ton/day sulfur-burning double absorption sulfuric acid plant (feed gas: 10% SO_2). *(Courtesy Lurgi GmbH, Frankfurt, Germany.)*

recovery requires higher operating temperatures for the absorbers and acid coolers. Venturi concurrent absorbers operating at acid temperatures of 130 to 140°C are installed in several plants. To recover acid heat directly as steam, it was necessary to increase the acid temperatures from 110 to 120°C to about 200°C.

In 1987 Monsanto Enviro-Chem demonstrated its new Heat Recovery System (HRS) at Namhae Chemical in South Korea.[36,37] The process is based on 310 stainless steel, which resists corrosion in 98.5 percent H_2SO_4 at temperatures up to 220°C. The intermediate absorber at Namhae takes 194°C gas from the converter third stage economizer and absorbs the SO_3 in 199°C, 99 percent acid. Recirculated acid from the absorber is cooled from 220°C in a 10-bar HRS boiler. The added energy recovery for this process is reported to increase the total energy recovery to 90 percent. Monsanto's proposed Monarch Process combines its HRW technology with the wet catalytic converter process (Lurgi) to increase heat recovery and shift it to high pressure stream production for electric power generation.[31]

Other Sources of Sulfuric Acid

Spent sulfuric acid usually is diluted in the process in which it is used: titanium dioxide pigment processing, plastics manufacture, and so on. The dilute acid may be used in processes requiring dilute acid or may be concentrated for reuse by a number of vacuum evaporation processes (Simonson-Mantius, Chemetics)[29] or by thin-film evaporation (DuPont, Bofors). In the submerged combustion distillation process, water is evaporated from the dilute acid by forcing hot flue gases from a fuel-fired burner below the acid surface (Chemico). Concentration of 75 percent acid to 95 to 98 percent H_2SO_4 by the Pauling-Plinke process is done by feeding the 75 percent acid to a stripping column fitted with a stirred cast iron pot mounted in a furnace. The acid concentration in the pot must be kept above 80 percent to minimize corrosion.

Chemetics has developed a process for treating spent alkylation sulfuric acid with nitric acid to produce a sulfuric acid that can be used to acidulate phosphate rock, the major use for sulfuric acid. The organic

contaminants are converted to carbon particles that are removed with the gypsum on filtration of the phosphoric acid. Special alloys are used in the fabrication of the acid reactor.

Production and Consumption of Sulfuric Acid

The world production of sulfuric acid (1989) is presented in Table 14.4.[26] Nearly two-thirds of the world's new sulfuric acid production derives from brimstone burning. The balance of new acid production comes largely from pyrite combustion and nonferrous ore smelting. Recycled spent sulfuric acid is a significant fraction of acid production only in the United States and Oceania, primarily Australia. In the United States 52 percent of the spent acid comes from petroleum refineries, 38 percent from plastic and synthetic materials producers, and the balance from a wide variety of sources.

New sulfuric acid production and the number of sulfuric acid plants in the United States during the 1980s are given in Table 14.5 for major acid-producing states. The South accounts for about 75 percent of new sulfuric acid production, most of which is captive consumption for phosphate fertilizer production. With its massive phosphate rock deposits, Florida leads the nation in sulfuric acid production, to feed its phosphoric acid plants. Sulfur mines and recovered sulfur from oil and gas operations in Texas and Louisiana provide brimstone feed to sulfuric acid plants. Many of these plants are sited near major waterways to provide cheap transportation for raw materials and fertilizer products. Southwestern sulfuric acid plants usually are associated with smelters that ship the acid by rail to fertilizer plants.

Table 14.6 lists the sulfuric acid–consuming industries in the United States and shows the trends in their acid consumption rates through the 1980s. Agriculture accounted for 70 percent of U.S. consumption of sulfuric acid in 1989. Nearly all of the sulfuric acid

TABLE 14.4 World Production of New Sulfuric Acid for 1989 (Thousands of Metric Tons, 100% H_2SO_4)

	Sulfur Source			
	Brimstone	*Pyrite*	*Other*	*Total*
West Europe	12,269	6,943	5,655	24,867
France	3,500	—	580	4,080
Germany FR	1,590	800	1,300	3,690
Spain	10	2,550	700	3,260
East Europe	24,522	5,872	8,141	38,535
Poland	2,500	—	550	3,100
Former Soviet Union	18,720	4,250	5,350	28,320
North America	34,462	450	5,551	40,463
Canada	1,200	450	1,910	3,560
U.S.	33,262	—	3,641	36,903
Latin America	7,630	272	2,268	9,900
Brazil	3,050	202	450	3,702
Mexico	3,373	—	792	4,165
Asia	12,724	10,284	7,649	30,657
China PR	950	8,758	1,700	6,881
India	3,625	20	260	3,905
Japan	1,995	533	4,353	6,881
Oceania	1,380	—	600	1,980
Africa	10,054	1,534	1,240	12,828
Morocco	4,135	—	—	4,135
S. Africa	1,125	1,325	700	3,150
Tunisia	3,250	—	—	3,250
World Total	102,771	25,355	31,104	159,230

Source: Freeport Sulphur Marketing Department estimates.

TABLE 14.5 New Sulfuric Acid Production in the United States (Thousands of Metric Tons, 100% H_2SO_4)

	Number of Plants			*Sulfuric Acid Production*		
	1989	*1984*	*1979*	*1989*	*1984*	*1979*
Arizona	5	8	7	1,454	1,098	1,119
California	10	10	10	527	844	989
Florida	14	12	16	17,551	16,331	14,637
Idaho	2	2	3	(1)	(1)	1,257
Illinois	8	9	15	(1)	759	907
Louisiana	11	10	11	5,085	5,774	5,666
New Jersey	8	8	10	377	501	1,034
New Mexico	4	4	4	(1)	649	875
Oklahoma + Texas	15	17	17	1,772	1,842	3,107
Virginia	5	7	7	(1)	510	536
Unidentified	60	62	78	10,136	8,484	7,499
Total	142	149	178	36,903	36,794	37,627

Source: Hoffman, Current Industrial Reports, Series MA28B, U.S. Dept. of Commerce, Bureau of the Census, 1989;[26] *World Sulphur and Sulphuric Acid Atlas*, The British Sulphur Corp. Ltd.[39]
1. Data withheld to avoid disclosing figures for individual companies; data included in total.

consumed in agriculture was reacted with phosphate rock (principally $Ca_9(PO_4)_6 \cdot CaF_2$) to produce phosphoric acid. Granular phosphate fertilizers are produced by ammoniating phosphoric acid to yield mono- and diammonium phosphates; ammonium sulfate also is produced as a fertilizer. Petroleum refining is the largest nonfertilizer use for U.S. sulfuric acid. The acid competes with hydrogen fluoride as a catalyst in petroleum alkylation reactions

TABLE 14.6 Sulfuric Acid-Consuming Industries in the United States (Thousands of Metric Tons, 100% H_2SO_4)

Consuming Industries	*1898*	*1984*	*1979*
Agriculture	26,909	26,751	24,075
Phosphatic	26,436	26,373	23,192
Nitrogenous	285	261	546
Pesticides	99	70	148
Other agricultural	79	47	189
Other Industries	11,491	11,083	14,062
Petroleum refining	2,088	2,067	2,407
Synthetic rubber and plastics	1,172	959	647
Rayon and cellulose	159	155	252
Other chemicals	2,731	3,162	4,183
Copper ore processing	1,797	1,043	2,119
Uranium and other ore processing	257	357	416
Iron and steel	181	270	880
Other metals	157	163	125
Pulp and paper	978	814	760
Pigments and paint	423	332	628
Other	1,548	1,761	1,591
Total	38,400	37,834	38,083

Source: Morse, "End Uses of Sulfur and Sulfuric Acid," Mineral Industry Surveys, U.S. Department of the Interior, Bureau of Mines, 1990.[32]

for gasoline production. Sulfuric acid acts as a catalyst in synthetic rubber and plastics manufacture. Copper ore leaching is carried out by distributing the acid over leach piles of the ore and collecting the copper-rich leachate for processing. Sulfuric acid from nearby smelters is used in copper, uranium, and vanadium ore leaching.

REFERENCES

1. Blicharz, M., "Different Methods of Sulphur Removal from Gas from Non-ferrous Metal Works," paper read at Sulphur 88, Vienna, Austria, Nov. 6–9, 1988.
2. British Sulphur Corp. Ltd., *Sulphur*, No. 205, 25–29 (Nov./Dec. 1989).
3. British Sulphur Corp. Ltd., *Sulphur*, No. 206, 22–33 (Jan./Feb. 1990).
4. British Sulphur Corp. Ltd., *Sulphur*, No. 193, 33–36 (Nov./Dec. 1987).
5. British Sulphur Corp. Ltd., *Sulphur*, No. 207, 21–25 (Mar./Apr. 1990).
6. British Sulphur Corp. Ltd., *Sulphur*, No. 207, 37–41 (Mar./Apr. 1990).
7. British Sulphur Corp. Ltd., *Sulphur*, No. 193, 26–30 (Nov./Dec. 1987).
8. British Sulphur Corp. Ltd., *Sulphur*, No. 203, 24–32 (July/Aug. 1990).
9. British Sulphur Corp. Ltd., *Sulphur*, No. 190, 16–19 (May/June 1987).
10. British Sulphur Corp. Ltd., *Sulphur*, No. 193, 21–22 (Nov./Dec. 1987).
11. British Sulphur Corp. Ltd., *Sulphur*, No. 190, 33–40 (May/June 1987).
12. British Sulphur Corp. Ltd., *Sulphur*, No. 188, 30–36 (Jan./Feb. 1987).
13. British Sulphur Corp. Ltd., *Sulphur*, No. 192, 30–39 (Sept./Oct. 1987).
14. British Sulphur Corp. Ltd., *Sulphur*, No. 191, 30–35 (Julu/Aug. 1987).
15. Davies, M., Hodgson, D. S., and Rodda, J. "Applications of SARAMET in H_2SO_4 Plants," paper read at Sulphur 88, Vienna, Austria, Nov. 6–9, 1988.
16. DeWolf, P., and Larison, E. L., *American Sulphuric Acid Practice*, McGraw-Hill, New York, 1921.
17. Douren, L., "Sandvik SX—The Future Material for Concentrated Sulfuric Acid," paper read at Sulphur 87, Houston, TX, Apr. 5–8, 1987.
18. Dudley, N., Barret, M., and Baldock, D. *The Acid Rain Controversy*, Earth Resources Research, London, 1985.
19. Duecker, W. W., and West, J. R. (Eds.), *The Manufacture of Sulfuric Acid*, Van Nostrand Reinhold, New York, 1959.
20. Elsworth, S. *Acid Rain*, Pluto Press, London, 1987.
21. Fairlie, A. M., *Sulfuric Acid Manufacture*, Reinhold, New York, 1936.
22. Goar, B. G., "Claus Plant Capacity Boosted by Oxygen-Enrichment Process," *Oil & Gas Journal*, pp. 39–41. (Sept. 30, 1985).
23. Haynes, W., *Brimstone—The Stone That Burns*, D. Van Nostrand Co., Princeton, NJ, 1959.
24. Heydorn, B., Lopez, D., and Lauriente, D., "Sulfur," in *Chemical Economics Handbook*, SRI International, Menlo Park, CA, 1987.
25. Heydorn, B., Aguiar, D., and Ferguson, A., "Sulfuric Acid," in *Chemical Economics Handbook*, SRI International, Menlo Park, CA, 1987.
26. Hoffman, G. B., "Inorganic Fertilizer Materials and Related Products," Current Industrial Reports MA28B, U.S. Dept. of Commerce, Bureau of the Census, Washington, DC, 1989.
27. Holdschick, H., and Jurascheck, M. A., "Advances in the Development of Cast Materials for High Concentrated Sulphuric Acid at Temperatures up to 280°C," paper read at Sulphur 88, Vienna, Austria, Nov. 6–9, 1988.
28. Hyne, J., "The Forming, Handling, and Transportation of Elemental Sulphur," paper read at Sulphur 1990, Cancun, Mexico, Apr. 1–4, 1990.
29. Lebel, G., "Pilot Plant Development of the Chemetics TiO_2 Waste Acid Recovery Process," paper read at Sulphur 1990, Cancun, Mexico, Apr. 1–4, 1990.
30. Matros, Y. S., and Bunimovich, G. A., "Reverse Process of SO_2 Oxidation in Sulfuric Acid Production," paper read at Sulphur 1990, Cancun, Mexico, Apr. 1–4, 1990.
31. McAlister, D. R., Grendel, R. W., Schneider, D. R., Shafer, J. R., and Tucker, J. S., "A Sulfuric Acid Plant for the 1990s," paper read at Sulphur 1990, Cancun, Mexico, Apr. 1–4, 1990.
32. Morse, D. E., "End Uses of Sulfur and Sulfuric Acid in 1989," Mineral Industry Surveys, U.S. Dept. of the Interior, Bureau of Mines, Washington, DC, 1990.
33. Morse, D. E., "Sulfur," *Minerals Yearbook*, U.S. Dept. of the Interior, Bureau of Mines, Washington, DC, 1990.
34. Sames, J. A., and Paskall, H. G., "Can Oxygen Enrichment Replace Tail Gas Clean-up in Sulphur Recovery?," paper read at Sulphur 87, Houston, TX, Apr. 5–8, 1987.
35. Sander, U. H. F., Fischer, H., Rothe, U., and Kola, R. *Sulphur, Sulphur Dioxide, and Sulphuric Acid*, Verlag Chemie, Weinheim, 1984.

36. Shafer, J. R., "Namhae Heat Recovery System Update," paper read at Fertilizer Industry Round Table, Baltimore, MD, Nov. 15, 1988.
37. Smith, R. M., Sheputis, J., Kim, U. B., and Chin, Y. B., "Sulfuric Acid Heat Recovery System (HRS) Operations at Namhae Chemical Corporation, Korea," paper read at Sulphur 88, Vienna, Austria, Nov. 6–9, 1988.
38. Sohn, H. Y., George, D. B., and Zunkel, A. D., *Advances in Sulfide Smelting*, The Metallurgical Society of AIME, Warrendale, PA, 1983.
39. The British Sulphur Corp. Ltd., *World Sulphur and Sulphuric Acid Atlas*, 5th ed., Purley Press Ltd., London, 1989.
40. Thieler, Erich, *Sulphur*. Theodor Steinkopff, Leipzig, 1936.
41. Thrash, L. A., "World Sulfur Derived from Oil, Gas Production," *Oil & Gas Journal*, 87–91. (July 9, 1990).
42. Wyld, Wilfried. 1923. *Sulphuric Acid and Sulphur Dioxide Raw Materials*. New York: D. Van Nostrand.
43. Friedman, L., 1988. "The Sulfuric Acid Plant Recent Advances—Reviews and Analysis," paper read at the AIChE Central Florida Meeting, Clearwater, FL, June 2, 1988.

15

Petroleum and Its Products

H. L. Hoffman*

THE NATURE OF PETROLEUM

Petroleum is a mixture of hydrocarbons—chemical combinations of hydrogen and carbon. When burned completely, the hydrocarbons should yield only water (H_2O) and carbon dioxide (CO_2). When the burning is incomplete, carbon monoxide (CO) and various oxygenated hydrocarbons are formed. Since most burning uses air, nitrogen compounds also exist. In addition, there are other elements associated with the hydrocarbons in petroleum such as sulfur, nickel, and vanadium, just to name a few.

Petroleum is found normally at great depth underground or below seabeds. It can exist as a gas, liquid, solid, or a combination of these three states. Drilling is used to reach the gaseous and liquid deposits of petroleum. Then they are brought to the surface through pipe. The gas usually flows under its own pressure. The liquid may flow from its own pressure or be forced to the surface by submerged pumps. Solid or semisolid petroleum is brought to the surface in a number of ways: by digging with conventional mining techniques, by gasifying or liquefying with high temperature stream, or by burning a portion of the material in the ground so that the remainder can flow to the surface.

Natural gas is the gaseous form of petroleum. It is mostly the single-carbon molecule, methane (CH_4). When natural gas is associated with liquid petroleum underground, the methane will come to the surface in admixture with some heavier hydrocarbons. The gas is then said to be a wet gas. These heavier hydrocarbons are isolated and purified in natural gas processing plants. The operation yields ethane (petrochemical feed), propane (LPG), butane (refinery blending stock), and hydrocarbon liquids (natural gas condensate).

When the underground natural gas is associated with solid hydrocarbons such as tar or coal, the methane will have little other hydrocarbons. Then the gas is said to be a dry gas.

Crude oil is the common name given to the liquid form of petroleum. In some writings, one will see reference to "petroleum and natural gas," suggesting petroleum and crude oil are used as synonymous terms. Some crude

*Hydrocarbon Processing, Houston, Texas.

oils have such great density that they are referred to as heavy oils and tars.

Tar sands are small particles of sandstone surrounded by an organic material called bitumen. The bitumen is so highly viscous and clings so tenaciously to the sandstone that it is easy to think of the mixture as a solid form of petroleum. Yet it is a mixture of high-density liquid on a supporting solid.

Oil shales are real petroleum solids. The curious thing about oil shales is that they do not contain petroleum crude oil. Instead, they contain an organic material called kerogen. The kerogen can be heated to yield a liquid substance called shale oil, which in turn can be refined into more conventional petroleum products.

Many products derived from petroleum are partly the consequence of the vast collection of hydrocarbons occurring in petroleum's natural state. A far more important factor is the ability of the hydrocarbon processing industries to transpose laboratory discoveries into large-scale commercial operations. Thus, the petroleum industry is an interesting study in applied organic chemistry and physical property manipulation.

Most of this discussion will concern the processing of petroleum crude oil, the most widely used form of petroleum resources. Natural gas processing will come in briefly at a few points. And since most of the world's petroleum is consumed as energy fuels, it is appropriate to begin with a brief review of the world's total energy situation.

Largest Energy Supplier

In Chapter 17, the point is made that coal offers a much more abundant primary source of energy than does petroleum. This is certainly true, but another fact remains: the world presently gets most of its energy from crude oil and natural gas. Petroleum is the major source of fuel used in transportation, manufacturing, and home heating.

Primary energy sources are defined as those coming from natural raw materials. Electricity is not included because it is a secondary energy source; that is, generated by consuming one or more of the other natural energy sources.

To put petroleum consumption into perspective, the primary energy sources considered here are: petroleum crude oil, natural gas, coal, hydropower (water to generate electricity), and nuclear energy. The quantities reported here will exclude energy from wood, peat, animal waste, and other sources—despite their importance to some localities. Documentation for these latter sources are sketchy, whereas the other energy sources are well documented.[1-4]

The common practice is to relate energy units to a common product—in this case, to petroleum liquid. For example, world consumption of crude oil and liquids (condensates) from natural gas in 1989 amounted to about 66 million barrels per day. If the amount of energy from other sources were converted to equivalent barrels of oil, the total world energy consumption in 1989 would be about 166 million daily barrels of oil equivalent. The relative distribution[1] of these sources is shown in Fig. 15.1. More energy (40%) comes from oil than from any other single source.
other single source.

Proven crude oil reserves determined for several earlier years[2] are shown in Table 15.1. Also included in this table are the annual production rates[3] and the ratio of the reserves to annual production.

Fig. 15.1. Total world primary energy consumption in 1989 was equivalent to about 166 million daily barrels of oil.

TABLE 15.1 Proven Crude Oil Reserves Compared With Production Rates

Year	Billions of Barrels: Proven Reserves, First of Year[2]	Billions of Barrels: Annual Production[3]	Ratio, Reserves Divided by Production, Years
1950	76	3.8	20
1955	154	5.6	27
1960	256	7.7	33
1965	342	11.0	31
1970	517	16.7	31
1975	569	19.3	29
1980	574	22.0	26
1985	696	19.6	36
1990*	933	21.8	43

*Author's estimate.

The ratio of reserves to production is a way of gauging natural reserves in terms of the years the proven reserves would sustain current production rates. Yet, if this measurement is taken literally, one might wonder how the petroleum industry grew and prospered during periods when there were less than 20 years of reserves. An explanation lies in the fact that the exploration and production segments of the petroleum industry seek and find oil sources to replace depleting ones. Note, for example, how the ratio jumped during the period between 1950 and 1955 as a result of the big oil reserves discovered in Middle Eastern countries.

Another interesting departure from past trends occurred in 1980, when oil consumption started to decline. The cause for the decline had its origins in 1974, when oil-exporting countries, especially countries in the Middle East, increased the price of their oil threefold. In 1980, the price of exported oil increased an additional twofold. These unusual price increases influenced the cost of other energy sources and set the stage for widespread energy conservation measures.

The oil price increases encouraged exploration for oil and other energy sources. At the same time, techniques were developed to recover more oil from existing reservoirs. Taken together, an increasing oil reserve and a slower energy growth rate placed the oil production rate in the early 1990s at a level that could be sustained by the then-known oil reserves for a longer time than in the recent past. But the continued use of oil as a major energy source will depend mostly on the ability to find more oil reserves.

Ultimate recoverable petroleum reserves have been estimated with wide variations, the consensus being between 1.5 to 2.5 trillion barrels (10^{12} bbl). Most estimators go to great lengths to explain the basis of their estimates. By comparison, proven reserves are in the neighborhood of 900 billion barrels. Proven reserves are generally taken to mean: "the oil remaining in the ground which geological and engineering information indicate with reasonable certainty to be recoverable in the future from known reservoirs under existing economic and operating conditions." [2]

Alternate feedstocks for refineries are being sought from coal, tar sands, and shale oil. Synthetic crude oil will be made from these raw materials so that conventional refining units can continue to be used to make consumer products.

In the meantime, top priority will be given to use crude oil to make liquid transportation fuels (because of their convenience) and petrochemical materials (because of their diversity of uses). It is a good bet that many of the gross conversion methods now applied to crude oil processing will be replaced in the future by more specific conversion. It is in this transition that knowledge of organic chemistry guides the development of new refining processes.

From Well to Refinery

A listing of crude oil production and consumption rates[1] for various countries and areas is shown in Table 15.2. Note that the countries in the Middle East produce much more oil than they consume, whereas the United States and countries in Western Europe consume much more than they produce. This situation points up the great importance of worldwide petroleum movement. The difference between production and

TABLE 15.2 Geographical Distribution of 1989 Production and Consumption[1] (millions of barrels daily)

Country/Area	*Produced*	*Consumed*
United States	9.2	17.3
Canada	2.0	1.8
Mexico	2.9	1.7
Central and S. America	4.2	3.6
Africa	6.1	2.0
W. Europe	4.2	12.9
USSR	12.1	8.8
Other E. Europe	0.3	1.8
Middle East	16.9	3.1
China	2.8	2.3
Japan	0.0	5.0
Far East and Oceania	3.7	5.6
World Total	64.4	65.9

Note: When consumption is more than production, the difference is made up from supplies held in storage.

consumption for any one period is balanced by oil being added to or removed from extensive storage areas around the world.

The growth of world refining capacity attempts to keep up with the growing demand for petroleum products. A measure of this growth[1] is shown in Fig. 15.2. The upper curve shows total refining capacity, while the lower curve shows the amount of crude oil run through the refineries.

One might wonder why refining capacity continued to surge ahead when crude throughput took a dip in the period from 1974 through 1976. For one thing, the amount of crude oil available for processing is subject to the whims of international trade. Since so much crude oil comes from some countries (notably, Middle Eastern countries) to be refined in other countries, international relations between countries strongly influence how much crude oil feedstock is available.

Another factor in refining growth is the time required to construct processing units. In highly industrialized countries such as the United States, Japan, and Western European countries there are mounting restrictions on new refinery sites. Thus, a decision to build a refinery and the actual completion of that refining capacity will take several year in order to fulfill local and governmental requirements. Then one to four years of actual construction activities are required before a new refinery will start processing feedstocks.

Refineries are located mostly in the countries consuming refined product. It is easier to transport crude oil to major refining centers than to transport separately the many individual products. The distribution of refining capacity[1] by areas for 1990 is depicted in Fig. 15.3.

The variety of ways crude oils are delivered to refineries is indicated by using U.S. refineries as an example. U.S. refineries get their feedstock via pipelines, tank trucks, barges, and oceangoing vessels.[4] The amount received by each of these routes is shown in Table 15.3. The large quantity coming by water explains why many refineries are located near oceans and why they own or lease such a large fleet of barges and oceangoing vessels.

Fig. 15.2. World refining crude oil throughput is about 80% of the total available refining capacity.

Fig. 15.3. Distribution of world refining capacity.

TABLE 15.3 Methods of Transportation for Crude Oil Received at U.S. Refineries, 1989[4]

Transportation Method	*Vol %*
Domestic crude oil	
Pipelines	40
Tankers and barges	14
Tank cars and trucks	2
Subtotal	56
Foreign crude oil	
Pipelines	15
Tankers and barges	29
Subtotal	44
Total receipts	100

PRODUCT NAMES

The distinction between refined products and petrochemicals is often a subtle one. In general, when the product is a fraction from crude oil that includes a fairly large group of hydrocarbons, the fraction is classified as a refined product. Examples of refined products are: gasoline, diesel fuel, heating oils, lubricants, waxes, asphalts, and petroleum coke.

By contrast, when the product from crude oil is limited to only one or two specific hydrocarbons of fairly high purity, the fraction is called a petrochemical. Examples of petrochemicals are: ethylene, propylene, benzene, toluene, and styrene—to name only a few.

There are many more identifiable petrochemical products than there are refined products. There are many specific hydrocarbons that can be derived from petroleum. However, these hydrocarbons lose individual identity when they are grouped into a refined product.

Refined Products

Most refined products at the consumer level are blends of several refinery streams. Product specifications determine which streams are suitable for a specific blend. Part of the difficulty of learning about refining lies in the industry's use of stream names that are different from the names of the consumer products.

Consider the listing in Table 15.4. The names in the last column should be familiar because they are used at the consumer level. Yet within a refinery, these products will be blended from portions of crude oil fractions having the names shown in the first column. To make matters worse, specifications and statistics for the industry are often reported under yet another set of names—those shown in the middle column of Table 15.4.

Gasoline at the consumer level, for example, may be called benzol or petrol, depending on the country where it is sold. In the early stages of crude oil processing, most gasoline components are called naphthas. Kerosine is

TABLE 15.4 Several Names for the Same Material

Crude oil cuts	*Refinery blends*	*Consumer products*
Gases	Still Gases	Fuel gas
	Propane/Butane	Liquefied petroleum gas (LPG)
Light/Heavy naphtha	Motor Fuel	Gasoline
	Aviation turbine, Jet-B	Jet fuel (naphtha type)
Kerosine	Aviation turbine, Jet-A	Jet fuel (kerosine type)
	No. 1 Fuel oil	Kerosine (range oil)
Light gas oil	Diesel	Auto and tractor diesel
	No. 2 fuel oil	Home heating oil
Heavy gas oil	No. 4 fuel oil	Commercial heating oil
	No. 5 fuel oil	Industrial heating oil
	Bright stock	Lubricants
Residuals	No. 6 fuel oil	Bunker C oil
	Heavy residual	Asphalt
	Coke	Coke

another example. It may be called coal oil to denote that it replaces stove oil (or range oil) once derived from coal. Kerosine's historical significance was first as an illuminating oil for lamps that once burned sperm oil taken from whales. But today, kerosine fractions go mostly into transportation fuels such as jet fuel and high quality No. 1 heating oil.

Product Specifications

Product application and customer acceptance set detailed specifications for various product properties. In the United States, the American Society for Testing and Materials (ASTM) and the American Petroleum Institute (API) are recognized for establishing specifications on both products and methods for testing. Other countries have similar referee organizations. For example, in the United Kingdom, it is the Institute of Petroleum (IP). In West Germany, it is Deutsches Institute suer Normung (DIN). In Japan, it is the Ministry of International Trade and Industry (MITI).

Boiling range is the major distinction among refined products, and many other properties are related directly to the products in these boiling ranges. A summary of ASTM specifications for fuel boiling ranges[5] is given in Table 15.5.

Boiling range also is used to identify individual refinery streams—as an example will show in a later section concerning crude oil distillation. The temperature that separates one fraction from an adjacent fraction will differ from refinery to refinery. Factors influencing the choice of cut point temperatures includes the following: type of crude oil feed, kind and size of downstream processes, and relative market demand among products.

Other specifications can involve either physical or chemical properties. Generally these specifications are stated as minimum or maximum quantities. Once a product qualifies to be in a certain group, it may receive a premium price by virtue of exceeding minimum specifications or by being below maximum specifications. Yet all too often, the only advantage for being better than specifications is an increase in the volume of sales in a competitive market.

The evolution of product specifications will, at times, appear sadly behind recent developments in more sophisticated analytical techniques. Certainly the ultimate specifications should be based on how well a product performs in use. Yet the industry has grown comfortable with certain comparisons, and these standards are retained for easier comparisons with earlier products. Thus, it is not uncommon to find petroleum products sold under an array of tests and specifications—some seemingly measuring similar properties.

It is behind the scenes that sophisticated analytical techniques prove their worth. These techniques are used to identify specific hydrocarbons responsible for one property or another. Then suitable refining processes are devised to accomplish a desired chemical reaction that will increase the production of specific types of hydrocarbons.

In the discussion on refining schemes, major specifications will be identified for each product category. It will be left to the reader to remember that a wide variety of other specifications also must be met.

Product Yields

As changes occur in relative demand for refined products, refiners turn their attention to ways that will alter internal refinery streams. The big problem here is that the increase in volume of one fraction of crude oil will deprive some other product of that same fraction. This point is often overlooked when the question arises: "How much of a specific product can a refinery make?" Such a question should always be followed by a second question: "What other products will be penalized?"

Envision, for example, what would happen if the refining industry were to make all the gasoline it possibly could with today's present technology. The result would be to rob many other petroleum products. A vehicle that needs gasoline for fuel also needs such products as industrial fuels to fabricate the

TABLE 15.5 Major Petroleum Products and Their Specified Boiling Range[5]

Product designation	*ASTM designation*	*Specified temp. for vol % distilled at 1 atm., °F* 10%	50%	90%
Liquefied petroleum gas (LPG)	D 1835			
Commercial propane		—[a]		—[b]
Commercial butane		—[a]		—[c]
Aviation gasoline (Avgas)	D 910	167 max	221 max	275 max[d]
Automotive gasoline	D 439			
Volatility class A		158 max	170–250	374 max[e]
Volatility class B		149 max	170–245	374 max[e]
Volatility class C		140 max	170–240	365 max[e]
Volatility class D		131 max	170–235	365 max[e]
Volatility class E		122 max	170–230	365 max[e]
Aviation turbine fuel	D 1655			
Jet A or A-1		400 max		—[f]
Jet B		—[g]	374 max	473 max
Diesel fuel oil	D 975			
Grade 1-D				550 max
Grade 2-D				540–640
Grade 4-D		—	not specified	—
Gas turbine fuel oil	D 2880			
No. 0-GT		—[h]	not specified	—
No. 1-GT				550 max
No. 2-GT				540–640
No. 3-GT		—	not specified	—
No. 4-GT		—	not specified	—
Fuel oil	D 396			
Grade No. 1		419 max		550 max
Grade No. 2		—[h]		540–640
Grade No. 4		—	not specified	—
Grade No. 5		—	not specified	—
Grade No. 6		—	not specified	—

[a]vapor pressure specified instead of front end distillation
[b]95% point, −37°F max
[c]95% point, 36°F max
[d]final point, 338°F max
[e]final point, all classes, 437°F max
[f]final point, 572°F max
[g]20% point, 293°F max
[h]flash point specified instead of front end distillation

vehicle, lubricants for the engine's operation, asphalt for roads upon which the vehicle is to move, and petrochemical plastics and fibers for the vehicle's interior. Until adequate substitutes are found for these other petroleum products, it would be unwise to make only one product, even though sufficient technology may exist to offer this option.

This is not to say that substitutes will not be found, that these substitutes will not be better than petroleum products. In fact, many forecasts suggest that petroleum will ultimately be allocated only to transportation fuels and petrochemical feedstocks. It appeaars that these uses are the most suitable options for petroleum crude oil.

In the United States, the relative portions of refined products[4] made from crude oil are shown in Table 15.6. This distribution of products is the result of a long-standing trend to convert the heavier less valuable fractions into lighter more valuable fractions. The ways

TABLE 15.6 Product Yields from U.S. Refineries, 1989[4]

Product	*Vol % of Refinery Input*
Still gas	4.9
Liquefied gas	4.0
Gasoline	45.9
Jet fuel	10.1
Kerosine	0.5
Special naphtha	0.4
Petrochemical feed	2.5
Distillates	20.8
Lubricants	1.2
Waxes	0.1
Coke	3.9
Asphalt/road oil	3.1
Residuals	6.9
Miscellaneous	0.5
Total	104.8*

*Includes 4.8 vol % gain because most products are of less density than original feedstock.

this can be done will be discussed in the section on refinery schemes.

Petrochemicals

The portion of crude oil going to petrochemicals may appear small compared to fuels, but the variety of petrochemicals is huge. The listing in Table 15.7 will give some idea of the range of petrochemical applications. Many of these products are described in Chapter 22. A few will be included here as they come into competition with the manufacture of fuels.

Despite their variety, all commercially manufactured petrochemicals account for the consumption of only a small part of the total crude oil processed. In the United States, where petrochemicals have grown swiftly, the total petrochemical output at the beginning of 1990 was little more than 7 volume percent (vol %) of all petroleum feedstocks. An estimated 2 vol % went to energy of conversion, leaving 5 vol % represented as petrochemical products. Even so, this quantity has been sufficient for petrochemical-based materials to replace many products once made from such raw materials as coal, lumber, metal ores, and so forth.

REFINING SCHEMES

A refinery is a massive network of vessels, equipment, and pipes. The total scheme can be divided into a number of unit processes. In the discussion to follow, only major flow streams will be shown, and each unit will be depicted by a single block on a simplified flow diagram. Details will be discussed later.

TABLE 15.7 Petrochemical Applications

Absorbents	De-emulsifiers	Hair conditioners	Pipe
Activators	Desiccants	Heat transfer fluids	Plasticizers
Adhesives	Detergents	Herbicides	Preservatives
Adsorbents	Drugs	Hoses	Refrigerants
Analgesics	Drying oils	Humectants	Resins
Anesthetics	Dyes	Inks	Rigid foams
Antifreezes	Elastomers	Insecticides	Rust inhib.
Antiknocks	Emulsifiers	Insulations	Safety glass
Beltings	Explosives	Lacquers	Scavengers
Biocides	Fertilizers	Laxatives	Stabilizers
Bleaches	Fibers	Odorants	Soldering flux
Catalysts	Films	Oxidation inhib.	Solvents
Chelating agents	Finish removers	Packagings	Surfactants
Cleaners	Fire-proofers	Paints	Sweeteners
Coatings	Flavors	Paper sizings	Synthetic rubber
Containers	Food supplements	Perfumes	Textile sizings
Corrosion inhib.	Fumigants	Pesticides	Tire cord
Cosmetics	Fungicides	Pharmaceuticals	
Cushions	Gaskets	Photographic chem.	

Refined products establish the order in which each refining unit will be introduced. Only one or two key product specifications are used to explain the purpose of each unit. Nevertheless, the reader is reminded that the choice from among several types of units and the size of these units are complicated economic decisions. The trade-offs among product types, quantity, and quality will be mentioned to the extent that they influence the choice of one kind of process unit over another.

Feedstock Identification

Each refinery has its own range of preferred crude oil feedstock for which a desired distribution of products is obtained. The crude oil usually is identified by its source country, underground reservoir, or some distinguishing physical or chemical property. The three most frequently specified properties are density, chemical characterization, and sulfur content.

API gravity is a contrived measure of density.[6] The relation of API gravity to specific gravity is given by the following:

$$°\text{API} = \frac{141.5}{sp\ gr} - 131.5$$

where *sp gr* is the specific gravity, or the ratio of the weight of a given volume of oil to the weight of the same volume of water at a standard temperature, usually 60°F.

An oil with a density the same as that of water, or with a specific gravity of 1.0, would then be 10°API oil. Oils with higher than 10°API gravity are lighter than water. Since lighter crude oil fractions are usually more valuable, a crude oil with a higher °API gravity will bring a premium price in the market place.

Heavier crude oils are getting renewed attention as supplies of lighter crude oil dwindle. In 1967, the U.S. Bureau of Mines (now part of the U.S. Department of Energy) defined heavy crudes as those of 25°API or less. More recently, the American Petroleum Institute proposed to use 20°API or less as the distinction for heavy crude oils.

A characterization factor was introduced by Watson and Nelson[7] to use as an index of the chemical character of a crude oil or its fractions. The Watson characterization factor is defined as follows:

$$\text{Watson } K = (T_B)^{1/3}/sp\ gr$$

where T_B is the absolute boiling point in degrees Rankin, and *sp gr* is specific gravity compared to water at 60°F. For a wide boiling range material like crude oil, the boiling point is taken as an average of the five temperatures at which 10, 30, 50, 70, and 90 percent are vaporized.

A highly paraffinic crude oil might have a characterization factor as high as 13, whereas a highly naphthenic crude oil could be as low as about 10.5. Highly paraffinic crude oils can also contain heavy waxes which make it difficult for the oil to flow. Thus, another test for paraffin content is to measure how cold a crude oil can be before it fails to flow under specific test conditions. The higher the pour point temperature, the greater the paraffin content for a given boiling range.

Sour and *sweet* are terms referring to a crude oil's approximate sulfur content. In early days, these terms designated smell. A crude oil with a high sulfur content usually contains hydrogen sulfide—the gas associated with rotten eggs. Then the crude oil was called sour. Without this disagreeable odor, the crude oil was judged sweet. Today, the distinction between sour and sweet is based on total sulfur content. A sour crude oil is one with more than 0.5 wt % sulfur, whereas a sweet crude oil has 0.5 wt % or less sulfur. It has been estimated that 58 percent of U.S. crude oil reserves are sour. More important, an estimated 81 percent of world crude oil reserves are sour.[8]

ASTM distillation is a test prescribed by the American Society for Testing and Materials to measure the volume percent distilled at various temperatures.[5] The results are often reported the other way around: the temperatures at which given volume percents vaporize.[9] These data indicate the quantity of conventional boiling range products occurring naturally in the crude oil. Analytical tests on

Fig. 15.4. Analysis of Light Arabian crude oil.

each fraction indicate the kind of processing that may be needed to make specification products. A plot of boiling point, sulfur content, and API gravity for fractions of Light Abrabian crude oil is shown in Fig. 15.4. This crude oil is among the ones most traded in the international crude oil market.

In effect, Fig. 15.4 shows that the material in the mid-volume range of Light Arabian crude oil has a boiling point of approximately 600°F, a liquid density of approximately 30°API and an approximate sulfur content of 1.0 wt %. These data are an average of eight samples of Light Arabian crude oil. More precise values would be obtained on a specific crude oil if the data were to be used in design work.

Since a refinery stream spans a fairly wide boiling range, the crude oil analysis data would be accumulated throughout that range to give fraction properties. The intent here is to show an example of the relation between volume distilled, boiling point, liquid density, and sulfur content.

Crude Oil Pretreatment

Crude oil comes from the ground admixed with a variety of substances: gases, water, and dirt (minerals). The technical literature devoted to petroleum refining often omits crude oil cleanup steps. It is likely presumed that the reader wishing to compare refining schemes will understand that the crude has already been through these cleanup steps. Yet cleanup is important if the crude oil is to be transported effectively and to be processed without causing fouling and corrosion. Cleanup takes place in two ways: field separation and crude desalting.

Field separation is the first attempt to remove the gases, water, and dirt that accompany crude oil coming from the ground. As the term implies, field separation is located in the field near the site of the oil wells. The field separator is often no more than a large vessel that gives a quieting zone to permit gravity separation of three phases: gases, crude oil, and water (with entrained dirt).

The crude oil is lighter than water but heavier than the gases. Therefore, crude oil appears within the field separator as a middle layer. The water is withdrawn from the bottom to be disposed of at the well site. The gases are withdrawn from the top to be piped to a natural gas processing plant or are pumped back into the oil well to maintain well pressure. The crude oil from the middle layer is pumped to a refinery or to storage awaiting transportation by other means.

Crude desalting is a water-washing operation performed at the refinery site to get additional crude oil cleanup.[10] The crude oil coming from field separators will continue to have some water and dirt entrained with it. Water washing removes much of the water-soluble minerals and entrained solids.

If these crude oil contaminants were not removed, they would cause operating problems during refinery processing. The solids (dirt and silt) would plug equipment. Some of the solids, being minerals, would dissociate at high temperature and corrode equipment. Still others would deactivate catalysts used in some refining processes.

Crude Oil Fractions

The importance of boiling range for petroleum products already has been discussed in connection with Table 15.5. The simplest form of refining would isolate crude oil into

Fig. 15.5. Separating desalted crude oil into fractions.

fractions having boiling ranges that would coincide with the temperature ranges for consumer products. Some treating steps might be added to remove or alter undesirable components, and a very small quantity of various chemical additives would be included to enhance final properties.

Crude oil distillation separates the desalted crude oil into fractions of differing boiling ranges. Instead of trying to match final product boiling ranges at this point, the fractions are defined by the number and kind of downstream processes.

The desalting and distillation units are depicted in Fig. 15.5 to show the usual fractions coming from crude oil distillation units. The discussion in the following paragraphs shows the relationships between some finished products and downstream processing steps.

GASOLINE

The light and heavy naphtha fractions from crude oil distillation are ultimately combined to make gasoline. The two streams are isolated early in the refining scheme so that each can be refined separately for optimum blending in order to achieve required specifications—of which only volatility, sulfur content, and octane number will be discussed.

Volatility

A gasoline's boiling range is important during its aspiration into the combustion chamber of a gasoline-powered engine. Vapor pressure, a function of the fuel's boiling range, is also important. Boiling range and vapor pressure are lumped into one concept, volatility.[11]

Lighter components in the gasoline blend are established as a compromise between two extremes: enough light components are needed to give adequate vaporization of the fuel–air mixture for easy engine starting in cold weather, but too much of the light components can cause the fuel to vaporize within the fuel pump and result in vapor lock.

Environmental studies suggest that hydrocarbons in the atmosphere near large cities are the result of evaporation of lighter components from the gasoline in automobiles. This evaporation is reduced by designing automobile to use closed fuel systems and fuel-injected engines. Then the concentration of the lighter components in the fuel can be reduced and is not so critical as it is for fuel-aspirated engines.

Heavier components are a trade-off between fuel volume and combustion chamber deposits. They extend the yield of gasoline that can be made from a given volume of crude oil, but they also contribute to combustion chamber deposits and spark plug fouling. Thus, an upper limit is set on gasoline's boiling range to give a clean-burning fuel.

Sulfur Content

Sulfur compounds are corrosive and foul-smelling. When burned in an engine, these compounds result in sulfur dioxide exhaust. Should the engine be equipped with a catalytic muffler, as is the case for many modern automobile engines, the sulfur is exhausted from the muffler as sulfur trioxide, or sulfuric acid mist.

Caustic wash or some other enhanced solvent washing technique is usually sufficient to remove sulfur from light naphtha. The sulfur compounds in light naphtha are mercaptans and organic sulfides that are removed readily by these washing processes.

Heavy naphtha is harder to desulfurize. The sulfur compounds are in greater concentration and are of more complicated molecular structure. A more severe desulfurization method is needed to break these structures and release the sulfur. One such process is hydrotreating.

Hydrotreating is a catalytic process that converts sulfur-containing hydrocarbons into low-sulfur liquids and hydrogen sulfide.[12] The process is operated under a hydrogen-rich blanket at elevated temperature and pressure. A separate supply of hydrogen is needed to compensate for the amount of hydrogen required to occupy the vacant hydrocarbon site once held by the sulfur. Also, hydrogen is consumed to convert the sulfur to hydrogen sulfide gas.

Nitrogen and oxygen compounds are also dissociated by hydrotreating. The beauty of the process is that molecules are split at the points where these contaminants are attached. For nitrogen and oxygen compounds, the products of hydrotreating are ammonia and water, respectively. Thus, the contaminants will appear in the offgases and are easily removed by conventional gas treating processes.

Octane Number

Another condition to keep gasoline engines running smoothly is that the fuel–air mixture start burning at a precise time in the combustion cycle. An electrical spark starts the ignition. The remainder of the fuel–air mix should be consumed by a flame front moving out from the initial spark.

Under some conditions, a portion of the fuel–air mixture will ignite spontaneously instead of waiting for the flame front from the carefully timed spark. The extra pressure pulses resulting from spontaneous combustion are usually audible above the normal sounds of a running engine and give rise to the phenomenon called knock. Some special attributes of the knocking phenomenon are called pinging and rumble. All of these forms of knock are undesirable because they waste some of the available power of an otherwise smooth-running engine.

Octane number is a measure of a fuel's ability to avoid knocking. The octane number of a gasoline is determined in a special single cylinder engine where various combustion conditions can be controlled.[5] The test engine is adjusted to give trace knock from the fuel to be rated. Then various mixtures of iso-octane (2,2,4-trimethyl pentane) and normal heptane are used to find the ratio of the two reference fuels that will give the same intensity of knock as that from the unknown fuel. Defining iso-octane as 100 octane number and normal heptane as 0 octane number, the volumetric percentage of iso-octane in heptane that matches knock from the unknown fuel is reported as the octane number of the fuel. For example, 90 vol % iso-octane and 10 vol % normal heptane establishes a 90 octane number reference fuel.

Two kinds of octane number ratings are specified, although other methods are often used for engine and fuel development. Both methods use the same reference fuels and

essentially the same test engine. Engine operating conditions are the difference. In one, called the *Research method*, the spark advance is fixed, the air inlet temperature is 125°F and engine speed is 600 rpm. The other, called the *Motor method*, uses variable spark timing, a higher mixture temperature (300°F), and a faster engine speed (900 rpm).

The more severe conditions of the Motor method have a greater influence on commercial blends than they do on the reference fuels. Thus, a Motor octane number of a commercial blend tends to be lower than the Research octane number. Recently, it has become the practice to label gasoline with an arithmetic avarage of both ratings, abbreviated (R + M)/2.

Catalytic reforming is the principal process for improving the octane number of a naphtha for gasoline blending.[13] The process gets its name from its ability to re-form or re-shape the molecular structure of a feedstock. The transformation that accounts for the improvement in octane number is the conversion of paraffins and naphthenes to aromatics. The aromatics have better octane numbers than their paraffin or naphthene homologs. The greater octane number increase for the heavier molecules explains why catalytic reforming is usually applied to the heavy naphtha fractions.

Catalysts for reforming typically contain platinum or a mixture of platinum and other metal promoters on a silica–alumina support. Only a small concentration of platinum is used, averaging about 0.4 wt %. The need to sustain catalyst activity and the expense of the platinum make it common practice to pretreat the reformer's feedstock to remove catalyst poisons.

Hydrotreating, already discussed, is an effective process to pretreat reforming feedstocks. The two processes go together well for another reason. The reformer is a net producer of hydrogen by virtue of its cyclization and dehydrogenation reactions. Thus, the reformer can supply the hydrogen needed by the hydrotreating reactions. A rough rule of thumb is that a catalytic reformer produces 800 to 1200 scf of hydrogen per barrel of feed, while the hydrotreater consumes about 100 to 200 scf/bbl for naphtha treating. The excess hydrogen is available for hydrotreating other fractions in separate hydrotreaters.

DISTILLATES

Jet fuel, kerosine (range oil), No. 1 fuel oil, No. 2 fuel oil, and diesel fuel are all popular distillate products coming from 400°F to 600°F fractions of crude oil. One grade of jet fuel uses the heavy naphtha fraction, but the kerosine fraction supplies the more popular heavier grade of jet fuel, with smaller amounts sold as burner fuel (range oil) or No. 1 heating oil.

Some heating oil (generally No. 2 heating oil) and diesel fuel are very similar and are sometimes substitutes for each other. The home heating oil is intended to be burned within a furnace for space heating. The diesel fuel is intended for compression-ignition engines.

Hydrotreating improves the properties of all these distillate products. The process not only reduces the sulfur content of the distillates to a low level but also hydrogenates unsaturated hydrocarbons so that they will not contribute to smoke and particulate emissions —whether the fuel is burned in a furnace or used in an engine.

RESIDUALS

Crude oil is seldom distilled at temperatures above about 650°F. At higher temperatures, coke will form and plug the lower section of the crude oil distillation tower. Therefore, the portion with a boiling point above 650°F is not vaporized—or at least not with the processing units introduced so far. This residual liquid is disposed of as industrial fuel oils, road oils, and so forth. The residual is sometimes called reduced crude because the lighter fractions have been removed.

PRODUCING MORE LIGHT PRODUCTS

The refining scheme evolved to this point is shown in Fig. 15.6. It is typical of a low-investment refinery designed to make

Fig. 15.6. Low-investment route to modern products.

products of modern quality. Yet the relative amounts of products are dictated by the boiling range of the crude oil feed. For Light Arabian crude oil reported earlier (see Fig. 15.4), all distillate fuel oils and lighter products (those boiling below 650°F) would comprise only about 55 vol % of the crude oil feed rate.

For industrialized areas where the principal demand is for transportation fuels or high-quality heating oils, a refining scheme of the type shown in Fig. 15.6 would need to dispose of almost half of the crude oil as low-quality, less desirable, residual products. Moreover, the price obtained for these residual products is not only much lower than revenues from lighter products but also lower than the cost of the original crude oil. Thus, there are economic incentives to convert much of the residual portions into lighter products of suitable properties.

Relative volumes of petroleum product deliveries in the United States and the portions existing in Light Arabian crude oil are compared by boiling ranges in Fig. 15.7. Note that 80 to 85 vol % of all U.S. petroleum products are lighter than the boiling temperature of 650°F compared to the 55 vol % existing in the crude oil. Furthermore, half of all U.S. products are gasoline and lighter

Fig. 15.7. Light Arabian crude oil compared to U.S. product deliveries.

distillates (boiling temperatures less than 400°F) compared to 29 vol % in the crude oil.

This comparison can appear unfair because an array of products obtained easily from one crude oil would be difficult to obtain from another crude oil feed. Total product deliveries in the United States come from a variety of different crude oils, processed in a variety of different refining schemes. But the comparison serves to emphasize a long-term trend in the refining industry—to convert heavy, less desirable fractions into lighter, more valuable products. The comparison also lays the foundation for the next group of processes to be discussed. The choice and the arrangement of processes hereafter are intended to depict

a breadth of refining technology, rather than suggest a commercial scheme for handling the example crude oil.

Cracking

These processes cause hydrocarbon molecules to break apart into two or more smaller molecules. Thermal cracking uses high temperature (above 650°F) and long residence time to accomplish the molecular split. Catalytic cracking accomplishes the split much faster and at lower temperatures because of the presence of a cracking catalyst.

Catalytic cracking involves not only some of the biggest units, with their large catalyst reactor-separators and regenerator, but it is also among the more profitable operations with its effective conversion of heavy feeds to light products. Gasoline from catalytic cracking has a higher octane number than thermally cracked gasoline. Yields include less gas and coke than thermal cracking; that is, more useful liquid products are made. The distribution of products between gasoline and heating oils can be varied by different choices for catalysts and operating conditions.

The best feeds for catalytic crackers are determined by a number of factors. The feed should be heavy enough to justify conversion. This usually sets a lower boiling point of about 650°F. The feed should not be so heavy that it contains undue amounts of metal-bearing compounds or carbon-forming materials. Either of these substances is more prevalent in heavier fractions and can cause the catalyst to lose activity more quickly.

Visbreaking is basically a mild, once-through thermal-cracking process. It is used to get just sufficient cracking of resid so that fuel oil specifications can be made. Although some gasoline and light distillates are made, this is not the purpose of the visbreaker.

Coking is another matter. It is a severe form of thermal cracking in which coke formation is tolerated to get additional lighter liquids from the heavier, dirtier fractions of crude oil. Here, the metals that would otherwise foul a catalytic process are laid down with the coke. The coke settles out in large coke drums that are removed from service frequently (about one a day) to have the coke cut out with high-pressure water lances. To make the process continuous, multiple coke drums are used so that some drums can be onstream while others are being unloaded.

Hydrocracking achieves cracking with a rugged catalyst to withstand resid contaminants and with a hydrogen atmosphere to minimize coking. Hydrocracking combines hydrotreating and catalytic-cracking goals, but a hydrocracker is much more expensive than either of the other two. The pressure is so high (up to 3000 psi) that very thick walled vessels must be used for reactors (up to 9 inches thick). The products from a hydrocracker will be clean (desulfurized, denitrified, and demetalized) and will contain isomerized hydrocarbons in greater amount than in conventional catalytic cracking. A significant part of the expense of operating a hydrocracker is for the hydrogen that it consumes.

Vacuum Distillation

Among the greater variety of products made from crude oil, some of the products (lubricating oils for example) have boiling ranges that exceed 650°F—the general vicinity where cracking would occur in atmospheric distillation. Thus, by using a second distillation unit under vacuum, the heavier parts of the crude oil can continue to be divided into specific products. Furthermore, some of the fractions distilled from vacuum units are better than atmospheric residue for cracking because the metal-bearing compounds and carbon-forming materials are more highly concentrated in the vacuum residue.

Reconstituting Gases

Cracking processes to convert heavy liquids to lighter liquids also make gases. Another way to make more liquid products is to combine gaseous hydrocarbons. A few small molecules of a gas can be combined to make one bigger molecular with fairly specific properties. Here, a gas separation unit is added to the refinery scheme to isolate the individual types of gases. When catalytic

cracking is also part of the refining scheme, there will be a greater supply of olefins—ethylene, propylene, and butylene. Two routes for reconstituting these gaseous olefins into gasoline blending stocks are described below.

Polymerization ties two or more olefins together to make polymer gasoline. The double bond in only one olefin is changed to a single bond during each link between two olefins. This means the product will still have a double bond. For gasoline, these polymer stocks are good for blending because olefins tend to have higher octane numbers than their paraffin homolog.

However, the olefinic nature of polymer gasoline can also be a drawback. During long storage in warmer climates, the olefins can continue to link up to form bigger molecules of gum and sludge. This effect, though, is seldom important when the gasoline goes through ordinary distribution systems.

Alkylation combines an olefin and iso-butane when gasoline is desired. The product is mostly isomers. If the olefin were butylene, the product would contain a high concentration of 2,2,4-trimethyl pentane. The reader is reminded that this is the standard compound that defines 100 on the octane number scale. Alkylates are high-quality gasoline-blending compounds, having good stability as well as high octane numbers.

Ether processes combine an alcohol with an iso-olefin. This is a recent addition to the gasoline manufacturing scheme. These processes were prompted by newer regulations requiring gasoline blends to contain some oxygenated compounds. When the alcohol is methanol and the iso-olefin is isobutylene, the product is methyl tertiary butyl ether (MTBE). When the alcohol is ethanol, and product is ethyl tertiary butyl ether (ETBE). When the alcohol is methanol and the iso-olefin is isoamylene, the product is tertiary amyl methyl ether (TAME).

A MODERN REFINERY

A refining scheme incorporating the processes discussed so far is shown in Fig. 15.8. The variations are quite numerous, though. Types of crude oil available, local product demands, and competitive quality goals are just a few of the factors considered to decide a specific scheme.

Many other processes play an important role in the final scheme. A partial list of these other processes would have the following goals: dewaxing lubricating oils, deoiling waxes, deasphalting heavy fractions, manufacturing specific compounds for gasoline blending (alcohols, ethers, etc.), and isolating specific fractions for use as petrochemical feedstocks.

Petrochemicals

It has already been mentioned that petrochemicals account for only a little more than 7 vol % of all petroleum feedstocks. Earlier, Table 15.7 gave the vast array of the applications of these petrochemicals. Olefins and aromatics make up a big part of the total.

Ethylene is one of the most important olefins. It is usually made by cracking gases—ethane, propane, butane or a mixture of these as might exist in a refinery's offgases. When gas feedstock is scarce or expensive, naphthas and even whole crude oil have been used in specially designed ethylene crackers. The heavier feeds also give significant quantities of higher molecular weight olefins and aromatics.

Aromatics, as were pointed out, are in high concentration in the product from a catalytic reformer. When aromatics are needed for petrochemical manufacture, they are extracted from the reformer's product using solvents such as glycols or sulfolane, to name two popular ones.

The mixed aromatics are called BTX as an abbreviation for benzene, toluene, and xylene. The first two are isolated by distillation, and the isomers of the third are separated by partial crystallization. Benzene is the starting material for styrene, phenol, and a number of fibers and plastics. Toluene is used to make a number of chemicals, but most of it is blended into gasoline. Xylene use depends on the isomer, *para*-xylene going into polyester

Fig. 15.8. High conversion refinery.

and *ortho*-xylene going into phthalic anhydride. Both are involved in a wide variety of consumer products.

PROCESS DETAILS

So far, refining units have been described as they relate to other units and to final product specifications. Now, typical flow diagrams[15] of some major processes will be presented to highlight individual features. In many cases, the specific design shown is an arbitrary choice from the work of several equally qualified designers.

Crude Desalting

Basically a water-washing process, the crude desalter must accomplish intimate mixing between the crude oil and water, then separate them sufficiently so that water will not enter subsequent crude-oil distillation heaters.

A typical flow diagram is shown in Fig. 15.9. The unrefined crude oil is heated to 100 to 300°F for suitable fluid properties. The

Fig. 15.9. Crude desalting. Includes: heater (1), mixing valve (2), and electrostatic water settler (3). (Hydrocarbon Processing, *69, no. 11, p. 86, Nov. 1990; copyright 1990 by Gulf Publishing Co. and used by permission of the copyright owner.*)

operating pressure is 40 psig or more. Elevated temperatures reduce oil viscosity for better mixing, and elevated pressure suppresses vaporization. The wash water can be added either before or after heating.

Mixing between the water and crude oil is assured by passing the mixture through a throttling valve or emulsifier orifice. Trace quantities of caustic, acid, or other chemicals are sometimes added to promote treating. Then the water-in-oil emulsion is introduced into a high voltage electrostatic field inside a gravity settler. The electrostatic field helps the water droplets to agglomerate for easier settling.

Salts, minerals, and other water-soluble impurities in the crude oil are carried off with the water discharged from the settler. Clean desalted crude oil flows from the top of the settler and is ready for subsequent refining.

Additional stages can be used in series to get additional reduction in salt content of the crude oil. Two stages are typical, but some installations use three stages. The increased investment cost for multiple stages is offset by reduced corrosion, plugging, and catalyst poisoning in downstream equipment by virtue of lower salt content.

Crude Distillation

Single or multiple distillation columns are used to separate crude oil into fractions determined by their boiling range. Common identification of these fractions was discussed in connection with Fig. 15.5, but should only be considered as a guide because a variety of refining schemes call for altering the type of separation made at this point.

A typical flow diagram of a two-stage crude oil distillation system is shown in Fig. 15.10. The crude oil is heated by exchange with various hot products coming from the system before it passes through a fired heater. The temperature of the crude oil entering the first column is 600 to 700°F, or high enough to vaporize the heavy gas oil and all lighter fractions.

Because light products must pass from the feed point up to their respective drawoff point, any intermediate stream will contain some of these lighter materials. Stream stripping (note the group of steam strippers beside the first column in Fig. 15.10) is a way to reintroduce these light materials back into the tower to continue their passage up through the column.

The bottom stream from the first fractionating column goes into a second column operated under vacuum. Steam jet ejectors are used to create the vacuum so that the absolute pressure can be as low as 30 to 40 mm Hg (about 0.7 psia). The vacuum permits hydrocarbons to be vaporized at temperatures much below their normal boiling points. Thus, fractions with normal boiling points above 650°F can be separated by vacuum distillation without causing thermal cracking.

Lately, a popular addition to a crude distillation system has been a preflash column ahead of the two stages shown in Fig. 15.10. The preflash tower strips out the lighter portions of a crude oil before the remainder enters the atmospheric column. It is the lighter portions that set the vapor loading in the atmospheric column, which, in turn, determines the diameter of the upper section of the column.

Fig. 15.10. Crude distillation—Foster Wheeler USA Corp. Includes: exchanger heaters (1, 3), desalter process (2), fired heater (4), main fractionator (5), overhead gas accumulator (6), sidestream strippers (7, 8, 9, 10), second fired heater (11), and vacuum fractionator (12). (Hydrocarbon Processing, *69, no. 11, p. 86, Nov. 1990; copyright 1990 by Gulf Publishing Co. and used by permission of the copyright owner.*)

Incidentally, total refining capacity of a facility is reported in terms of its crude-oil handling capacity. Thus, the size of the first distillation column, whether a preflash or an atmospheric distillation column, sets the reported size of the entire refinery. Ratings in barrels per stream day (bpsd) will be greater than barrels per calendar day (bpcd). Processing units must be shut down on occasion for maintenance, repairs, and equipment replacement. The ratio of operating days to total days (or bpcd divided by bpsd) is called an "onstream factor" or "operating factor." The ratio will be expressed as either a percent or a decimal. For example, if a refinery unit undergoes one shutdown period of one month during a three year duration, its operating factor is $(36 - 1)/36$, or 0.972, or 97.2 percent.

Outside the United States, refining capacity is given normally in metric tons per year. Precise conversion from one unit of measure to the other depends upon the specific gravity of the crude oil, but the approximate relation is one barrel per day equals 50 tons/year.

Hydrotreating

This is a catalytic hydrogenation process that reduces the concentration of sulfur, nitrogen, oxygen, metals, and other contaminants in a hydrocarbon feed. In more severe forms, hydrotreating saturates olefins and aromatics.

A typical flow diagram is shown in Fig. 15.11. The feed is pumped to operating pressure and mixed with a hydrogen-rich gas, either before or after being heated to the proper reactor inlet temperature. The heated mixture passes through a fixed bed of catalyst where exothermic hydrogenation reactions occur. The effluent from the reactor is then cooled and sent through two separation stages. In the first, the high pressure separator, unreacted hydrogen is taken overhead to be scrubbed for hydrogen sulfide removal; the cleaned hydrogen is then recycled. In the second, the lower-pressure separator takes off the remaining gases and light hydrocarbons from the liquid product. If the feed is

Fig. 15.11. Hydrotreating—Chevron Research and Technology Co. Includes: reactor (1), hot high pressure separator (2), hot low pressure separator (3), cold high pressure separator (4), cold low pressure separator (5), and product fractionator (6). (Hydrocarbon Processing, *69, no. 11, p. 114, Nov. 1990; copyright 1990 by Gulf Publishing Co. and used by permission of the copyright owner.*)

a wide-boiling range material from which several blending stocks are to be made, the hot low pressure separator is followed by a fractionation column to remove various liquid fractions.

The feed for hydrotreating can be a variety of different boiling range materials extending from light naphtha to vacuum residues. Generally, each fraction is treated separately to permit optimum conditions—the higher boiling materials requiring more severe treating conditions. For example, naphtha hydrotreating can be carried out at 200 to 500 psia and at 500 to 650°F, with a hydrogen consumption of 10 to 50 scf/bbl of feed. On the other hand, a residue hydrotreating process can operate at 1000 to 2000 psia and at 650 to 800°F, with a hydrogen consumption of 600 to 1200 scf/bbl.[14] Nevertheless, hydrotreating is such a desirable cleanup step that it can justify its own hydrogen manufacturing facilities, although the hydrogen-rich stream obtained as a by-product from catalytic reforming usually is sufficient for most operations.

Catalyst formulations constitute a significant difference among hydrotreating processes. Each catalyst is designed to be best suited to one type of feed or one type of treating goal.[16] When hydrotreating is done for sulfur removal, the process is called hydrodesulfurization, and the catalyst generally is cobalt and molybdenum oxides on alumina. A catalyst of nickel-molybdenum compounds on alumina can be used for denitrogenation and cracked-stock saturation.

Catalytic Reforming

Some confusion comes from the literature when the term "naphtha reforming" is used to designate processes to make synthesis gas—a mixture containing predominantly carbon monoxide and hydrogen. However, naphtha reforming has another meaning, which is the one intended here—production of an aromatic-rich liquid for use in gasoline blending.

A typical flow diagram is shown in Fig. 15.12. The feed is pumped to operating pressure and mixed with a hydrogen-rich gas before heating to reaction temperatures. Actually, hydrogen is a by-product of the dehydrogenation and cyclization reactions,

Fig. 15.12. Catalytic reforming—Engelhard Corp. Includes: multibed reactors (1, 2, 3, 4), common heater, hydrogen separator (5), and compressor (6). (Hydrocarbon Processing, *69, no. 11, p. 120, Nov. 1990; copyright 1990 by Gulf Publishing Co. and used by permission of the copyright owner.*)

but by sustaining a hydrogen atmosphere, cracking and coke formation are minimized.

The feed for catalytic reforming is mostly in the boiling range of gasoline to start with. The intent is to convert the paraffin and naphthene portions to aromatics. As an example, a 180 to 310°F fraction of Light Arabian crude oil was reported to have 8 vol % aromatics before catalytic reforming, but was 68 vol % aromatics afterwards. The feed paraffin content (69 vol %) was reduced to less than half, and the feed naphthene content (23 vol %) was almost completely absent in the product.[13]

The extent of octane number change with changes in molecular configuration is shown in Table 15.8, where normal paraffins and naphthenes are compared with their aromatic homologs.

If the naphthenes are condensed (multirings or indanes) they tend to deactivate the reforming catalyst quickly. Control of the end point of the feed will exclude these deactivating compounds.

TABLE 15.8 Aromatics Have Higher Octane Numbers[6]

Hydrocarbon homologs		*Octane number, clear* Motor	Research
C_7 hydrocarbons			
n-paraffin	C_7H_{16} (*n*-heptane)	0.0	0.0
naphthene	C_7H_{14} (cycloheptane)	40.2	38.8
	" (methylcyclohexane)	71.1	74.8
aromatic	C_7H_8 (toluene)	103.5	120.1
C_8 hydrocarbons			
n-paraffin	C_8H_{18} (*n*-octane)	−15[a]	−19[a]
naphthene	C_8H_{16} (cyclooctane)	58.2	71.0
	" (ethylcyclohexane)	40.8	45.6
aromatic	C_8H_{10} (ethylbenzene)	97.9	10.4
	" (*o*-xylene)	100.0	120[a]
	" (*m*-xylene)	115.0	117.5
	" (*p*-xylene)	109.6	116.4

[a]Blending value at 20 vol % in 60 octane number reference fuel.

TABLE 15.9 Favored Operating Conditions for Desired Reaction Rates[17]

Feed	*Reaction*	*Product*	*Desired rate*	*To get desired rate* *Press.*	*Temp.*
Paraffins	Isomerization	Iso-paraffins	Inc.	Inc.	Inc.
	Dehydrocyclization	Naphthenes	Inc.	Dec.	Inc.
	Hydrocracking	Lower mol. wt.	Dec.	Dec.	Dec.
Naphthenes	Dehydrogenation	Aromatics	Inc.	Dec.	Inc.
	Isomerization	Iso-paraffins	Inc.	Inc.	Inc.
	Hydrocracking	Lower mol. wt.	Dec.	Dec.	Inc.
Aromatics	Hydrodealkylation	Lower mol. wt.	Dec.	Dec.	Dec.

Catalysts that promote reforming reactions can give side-reactions. Isomerization is acceptable, but hydrocracking gives unwanted saturates and gases. Therefore, higher operating pressures are used to suppress hydrocracking. This remedy has disadvantages. Higher pressures suppress reforming reactions too, although to a lesser extent. Generally, a compromise is made between desired reforming and undesired hydrocracking. The effects of operating conditions on competing reactions[17] are shown in Table 15.9.

In the late 1960s, it was discovered that the addition of certain promoters, such as rhenium, germanium, or tin, to the platinum-containing catalyst would reduce cracking and coke formation. The resulting catalysts, referred to as bimetallic catalysts, permit the process to enjoy the better reforming conditions of lower pressure without being unduly penalized by hydrocracking. Earlier pressures of 500 psig are now down to 150 psig.

Operating temperatures are important, too. The reactions are endothermic. Best yields would come from isothermal reaction zones, but they are difficult to achieve. Instead, the reaction beds are separated into a number of adiabatic zones operating at 500 to 1000°F with heaters between stages to supply the necessary heat of reaction and hold the overall train near a constant temperature. Three or four reactor zones are commonly used when it is desired to have a product with high octane numbers.

In making gasoline with high octane numbers but without the use of antiknock additives, high severity catalytic reforming is the prime route. The big disadvantage is a yield loss. Newer catalysts make the loss less dramatic, but the penalty remains,[18] as can be seen from Fig. 15.13.

Fig. 15.13. Better octane numbers, less yield (Kuwait naphtha). (*McDonald, G. W. G.,* Hydrocarbon Processing, *56 (6), 147–150, June 1977; copyright 1977 by Gulf Publishing Co. and used by permission of the copyright owner.*)

Catalytic Cracking

A typical diagram of a fluid catalytic cracking is shown in Fig. 15.14. The unit is characterized by two huge vessels, one to react the feed with hot catalyst and the other to regenerate the

Fig. 15.14. Fluid catalytic cracking—UOP. Includes: light recycle gas diluent addition at base of reactor (1), preacceleration zone (2), feed addition distributor (3), catalyst separator (4), catalyst regenerator (5), and catalyst cooler (6). (Hydrocarbon Processing, *69, no. 11, p. 98, Nov. 1990; copyright 1990 by Gulf Publishing Co. and used by permission of the copyright owner.*)

spent catalyst by burning off carbon with air. The activity of molecular-sieve catalysts is so great that the contact time between feed and catalyst is reduced drastically. If not, the oil will overcrack to give unwanted gases and coke. The short contact time is accomplished by using a transfer line between the regenerator and reactor vessels. In fact, the major portion of the reaction occurs in this piece of pipe or riser, and the products are taken quickly overhead. The main reactor vessels then are used to hold cyclone separators to remove the catalyst from the vapor products and to give additional space for cracking the heavier portions of the feed.

There are several configurations of reactors and regenerators. In some designs, one vessel is stacked on top of the other. All are big structures (150–200 ft high).

Riser cracking, as the short-time contacting is called, has a number of advantages. It is easier to design and operate. It can be operated at higher temperatures to give more gasoline olefins. It minimizes the destruction of any aromatics formed during cracking. The net effect can be the production of gasoline having octane numbers two or three numbers higher than earlier designs would give.

Better regeneration of the spent catalyst is obtained by operating at higher temperatures (1300–1400°F).[19] The coke that is deposited on the catalyst is more completely burned away by higher-temperature air blowing. The newer catalysts are rugged enough to withstand the extra heat, and newer metallurgy gives the regenerator vessel the strength it needs at higher temperatures.

Heavier feedstocks can be put into catalytic crackers. The nickel, vanadium, and iron in these heavier fractions do not deactivate the catalysts as fast as they once did because passivators are available now to add to the catalysts.[20] The extra sulfur that comes with heavier feeds can be prevented from exhausting into the atmosphere during regeneration because of catalysts that hold onto the sulfur compounds until the catalysts get into the reactor.[21] Then the sulfur compounds are cracked to light gases and leave the unit with the cracked products. Ordinary gas treating methods are used to capture the hydrogen sulfide coming from the sulfur in the feedstock.

Coking

Coking is an extreme form of thermal cracking. The process converts residual materials that might not easily be converted by the more popular catalytic cracking process. Coking is also a less expensive process for getting more light stocks from residual fractions. In the coking process, the coke is considered a by-product that is tolerated in the interest of more complete conversion of residues to lighter liquids.

A typical flow diagram of a delayed coker is shown in Fig. 15.15. There are several possible configurations, but in this one the feed goes directly into the product fractionator to pick up heavier products to be recycled to the cracking operation. The term "delayed coker" signifies that the heat of cracking is added by the furnace, and the cracking occurs during the longer residence time in the following coke drums. Furnace outlet temperatures are in the range of 900 to 950°F while the coke drum pressures are in the range of 15 to 90 psig.

The coke accumulates in the coke drum, and the remaining products go overhead as vapors to be fractionated into various products. In this case, the products are gas, naphtha, light gas oil, heavy gas oil, and coke. When a coke drum is to be emptied, a large drilling structure mounted on top of the drum is used to make a center hole in the coke formation. The drill is equipped with high-pressure water jets (3000 psig or more) to cut the coke from the drum so that it can fall out a bottom hatch into a coke pit. From there, belt conveyors and bucket cranes move the coke to storage or to market.[22]

Fluid coking is a proprietary name given to a different type of coking process in which the coke is suspended as particles in fluids flowing from a reactor to a heater and back again. When part of the coke is gasified, the process is called *Flexicoking*. Both Fluid coking and Flexicoking are proprietary processes of Exxon Research and Engineering Co.

A flow diagram for Flexicoking is shown in Fig. 15.16. The first two vessels are typical of Fluid coking, in which part of the coke is burned in the heater in order to have hot coke nuclei to contact the feed in the reactor vessel. The cracked products are quenched in an overhead scrubber where entrained coke is

Fig. 15.15. Delayed coking—ABB Lummus Crest Inc. Includes: feed/product fractionator (1), heater (2), coke drums (3), and vapor recovery unit (4). (Hydrocarbon Processing, *69, no. 11, p. 106, Nov. 1990; copyright 1990 by Gulf Publishing Co. and used by permission of the copyright owner.*)

Fig. 15.16. Fluid coking (Flexicoking)—Exxon Research and Engineering Co. Includes: reactor (1), scrubber (2), heater (3), gasifier (4), coke fines removal (5), and hydrogen sulfide removal (6). (Hydrocarbon Processing, *69, no. 11, p. 106, Nov. 1990; copyright 1990 by Gulf Publishing Co. and used by permission of the copyright owner.*)

returned to the reactor. Coke from the reactor circulates to the heater where it is devolatilized to yield a light hydrocarbon gas and residual coke. A sidestream of coke is circulated to the gasifier, where, for most feedstocks, 95 percent or more of the gross coke product from the reactor is gasified at elevated temperature with steam and air. Sulfur that enters the unit with the feedstock eventually becomes hydrogen sulfide exiting the gasifier and is recovered by a sulfur removal step.

Hydrocracking

Before the late 1960s, most hydrogen used in processing crude oil was for pretreating catalytic reformer feed naphtha and for desulfurizing middle-distillate products. Soon thereafter, requirements to lower sulfur content in most fuels became an important consideration. The heavier fractions of crude oil were the hardest to treat. Moreover, these fractions were the ones offering additional sources of light products. This situation set the stage for the introduction of hydrocracking.

A typical flow diagram for hydrocracking is shown in Fig. 15.17. Process flow is similar to hydrotreating in that feed is pumped to operating pressure, mixed with a hydrogen-rich gas, heated, passed through a catalytic reactor, and distributed among various fractions. Yet the hydrocracking process is unlike hydrotreating in several important ways. Operating pressures are very high, 2000 to 3000 psia. Hydrogen consumption also is high, 1200 to 1600 scf of hydrogen per barrel of feed, depending on the extent of the cracking.[14] In fact, it is not uncommon to see hydrocrackers built with their own hydrogen manufacturing facilities nearby.

The catalysts for hydrocracking have a dual function. They give both hydrogenation and dehydrogenation reactions and have a highly acidic support to foster cracking. The hydrogenation–dehydrogenation components of the catalysts are metals such as cobalt, nickel, tungsten, vanadium, molybdenum, platinum, palladium, or a combination of these metals. The acidic support can be silica–alumina, silica–zirconia, silica–magnesia, alumina–boria, silica–titania, acid-treated clays, acidic-metal phosphates, or alumina, to name some given in the literature.[23]

Great flexibility is attributed to most hydrocracking processes. Under mild conditions, the process can function as a hydrotreater. Under more severe conditions of cracking, the process produces a varying ratio of motor fuels and middle distillates, depending

Fig. 15.17. Hydrocracking—UOP. Includes: staged reactors (1, 2), gas separator (3), hydrogen separator (4), and product washer (5). (Hydrocarbon Processing, *69, no. 11, p. 102, Nov. 1990; copyright 1990 by Gulf Publishing Co. and used by permission of the copyright owner.*)

on the feedstock and operating variables. Even greater flexibility is possible for the process during design stages when it can be tailored to change naphthas into liquefied petroleum gases or convert heavy residues into lighter products.

Because the hydrocracker is viewed as both a cracker and a treater, it can appear in refining process schemes in a number of different places. As a cracker, it is used to convert feeds that are too heavy or too contaminant-laden to go to catalytic cracking. As a treater, it is used to handle heating-oil fractions that need to be saturated to give good burning quality. But it is the trend to heavier feeds and lighter high-quality fuels that causes hydrocracking to offer advantages to future refining, even though the hydrocracking units are much more expensive to build and to operate.

The principle of an ebulliating catalyst bed is embodied in some proprietary designs, in contrast with the fixed-catalyst beds used in other versions of hydrocracking. The H-Oil process of Hydrocarbon Research, Inc. and the LC-Fining process jointly licensed by ABB Lummus Crest Inc., Oxy Research & Development Co., and Amoco Corp. are examples of hydrocracking processes that use an ebullient bed instead of a fixed bed of catalyst.

Polymerization

This process usually is associated with the manufacture of plastic films and fibers from light hydrocarbon olefins, with products such as polyethylene and polypropylene. As a gasoline manufacturing process, the polymerization of light olefins emphasizes a combination of only two or three molecules so that the resulting liquid will be in the gasoline boiling range.

For early polymerization units, the catalyst was phosphoric acid on a quartz or kieselguhr support. Many of these units were shut down when the demand for gasoline with increased octane numbers prompted the diversion of the olefin feeds to alkylation units that gave higher octane number products. Yet some refinery balances have more propylene than alkylation can handle, so a newer version of polymerization was introduced.[24] It is the Dimersol process of the Institut Français du Pétrole, for which the flow diagram is shown in Fig. 15.18.

Fig. 15.18. Polymerization (Dimersol)—Institute Français du Pétrole. Includes: a reactor (1) of several in series, catalyst removal using a caustic wash (2), water wash (3), and distillation (4). (Hydrocarbon Processing, *69, no. 11, p. 134, Nov. 1990; copyright 1990 by Gulf Publishing Co. and used by permission of the copyright owner.*)

The Dimersol process uses a soluble catalytic complex injected into the feed before it enters the reactor. The heat of reaction is taken away by circulating a portion of the bottoms back to the reactor after passing it through a cooling water exchanger. The product goes through a neutralizing system that uses caustic to destroy the catalyst so that the resulting polymer is clean and stable. Typical octane number ratings for the product are 81 Motor and 96.5 Research, unleaded.

Alkylation

This is another process that increases the total yield of gasoline by combining some of the gaseous light hydrocarbons to form bigger molecules boiling in the gasoline range. Alkylation combines isobutane with a light olefin, typically propylene and butylene. A flow diagram for an alkylation unit using sulfuric acid as a catalyst is shown in Fig. 15.19.

Common catalysts for gasoline alkylation are hydrofluoric acid or sulfuric acid. The reaction is favored by higher temperatures, but competing reactions among the olefins to give polymers prevent high-quality yields. Thus, alkylation usually is carried out at low temperatures in order to make the alkylation reaction predominate over the polymerization reactions. Temperatures for hydrofluoric acid–catalyzed reactions are approximately 100°F, and for sulfuric acid they are approximately

Fig. 15.19. Alkylation—Stratco, Inc. Includes: contactor (1), acid settler (2), flash drum (3), deisobutanizer (4), compressor (5), and depropanizer (6). (Hydrocarbon Processing, *69, no. 11, p. 132, Nov. 1990; copyright 1990 by Gulf Publishing Co. and used by permission of the copyright owner.*)

TABLE 15.10 Typical Alkylate Octane Numbers[25]

	C_2	C_3	Feed olefin $C_3 + C_4$	C_4
Research octane number, clear	101.5	90.5	93	96.5
Motor octane number, clear	93	89	91	95.5

50°F. Since the sulfuric acid–catalyzed reactions are carried out below normal atmospheric temperatures, refrigeration facilities are included.

Alkylate product has a high concentration of 2,2,4-trimethyl pentane, the standard for the 100 rating of the octane number scale. Other compounds in the alkylate are higher or lower in octane number, but the lower octane number materials predominate so that alkylate has a Research octane number in the range of 92 to 99. Developments are under way to slant the reactions in favor of the higher-octane materials.[25] Random samples of alkylate quality reported in the literature[26] are summarized in Table 15.10.

Ether Processes

Cooperative studies between automobile manufacturers and gasoline producers established the relationship of some gasoline components to automobile engine emissions. It has been shown that the use of gasolines containing some oxygenated compounds, such as alcohols or ethers, will cause the gasoline-fueled engine to emit less carbon monoxide.[27]

The favored alcohols are low molecular weight ones, methanol and ethanol. Commercial technology exists for making either alcohol from a variety of raw materials; namely, natural gas, petroleum crude oil, coal, or agricultural grain. A higher molecular weight alcohol, tertiary butyl alcohol (TBA), is another suitable gasoline blending compound. The major drawback in using alcohols in gasolines is that it creates a new set of conditions for safe fuel handling and efficient engine design. On the other hand, the alcohols could be used to make ethers. These have been shown to have fewer blending problems than alcohols and were more compatible with existing gasoline blends.

Another environmentally imposed restriction on gasoline blends was to reduce vapor pressure. The restriction was imposed to control hydrocarbon evaporative losses during gasoline storage and automobile refueling. The consequence of this restriction was to reduce the amount of C_4 hydrocarbons included in the gasoline blends.

Fig. 15.20. Ethers—Huels AG and UOP. Includes: guard bed (1), reactors (2, 3), butene column (4), and water wash (5, 6). (Hydrocarbon Processing, *69, no. 11, p. 128, Nov. 1990; copyright 1990 by Gulf Publishing Co. and used by permission of the copyright owner.*)

A common second feed to the ether manufacturing process, in addition to an alcohol, is isobutylene. Isobutylene is not only one form of the C_4 hydrocarbons, but it can be made from the other C_4 homologs. So not only is ether manufacture a way to make a desirable oxygenated compound for gasoline blending, but the ether process gives a way to use the C_4 hydrocarbons to enhance the antiknock quality of the gasoline without causing the blend to exceed the vapor pressure restriction.

Methyl tertiary butyl ether (MTBE) is the more popular ether blending compound. It is made by reacting methanol with isobutylene. A flow diagram of a typical process is shown in Fig. 15.20.

Other desirable ether compounds for gasoline blending can be made in a similar process. For example, ethyl tertiary butyl ether (ETBE) is made from ethanol and isobutylene. Tertiary amyl methyl ether (TAME) is made from methanol and isoamylene.

FUTURE PROCESSING

The manufacture of fibers, films, construction materials, and many synthetic organic chemicals made from petroleum is evolving at such a rapid rate that these subjects are covered in other, separate chapters of this book. Yet the greatest use for petroleum and its products now is to furnish fuels for heat and mobility.

There is an ever-changing economic balance between the need for energy fuels and that for other petroleum-derived products. Many decades have been spent in perfecting technology to give the least expensive fuels for the most efficient energy consumption. Future processing technology will add another dimension: increased concern about the relation of increased energy use to environmental changes. We are only beginning to identify and to quantify how future energy needs might be satisfied in environmentally compatible ways.

REFERENCES

1. U.S. Energy Information Administration, "International Energy Annual 1989," DOE/EIA-0219(89), pp. xiii–xvi (Feb. 1991).
2. "45th Annual International Outlook," *World Oil*, **211** (2), 25–26 (Aug. 1990); and earlier annual editions.
3. U.S. Dept. of the Interior, *Minerals Yearbook, Metals, Minerals and Fuels*, annually.
4. U.S. Energy Information Administration, "Petroleum Supply Annual 1989, Vol. 1." DOE/EIA-0340(89)/1, Table 44, p. 112 and Table 19, p. 44 (1990).
5. American Society for Testing and Materials, *1990 Annual Book of ASTM Standards*, Vol. 05.01 and 05.02, 1990.
6. American Petroleum Institute, *Technical Data Book, Petroleum Refining*, 4th ed., Publ. No. 999 (1983).
7. Watson, K. M., and Nelson, E. F., "Improved Methods for Approximating Critical and Thermal Properties of Petroleum Fractions," *Industrial and Engineering Chemistry*, **25**, 880 (1933).
8. Hoffman, H. L., "Sour Crude Limits Refining Output," *Hydrocarbon Processing*, **52** (9), 107–110 (Sept. 1973).
9. Ferrero, E. P., and Nichols, D. T., "Analyses of 169 Crude Oils From 122 Foreign Oilfields," U.S. Department of the Interior, Bureau of Mines, Information Circular 8542 (1972).
10. Waterman, L. C., "Crude Desalting: Why and How," *Hydrocarbon Processing*, **44** (2), 133–138 (Feb. 1965).
11. Unzelman, G. H., and Forster, E. J., "How to Blend for Volatility," *Petroleum Refiner*, **39** (9), 109–140 (Sept. 1960).
12. Kay, H., "What Hydrogen Treating Can Do," *Petroleum Refiner*, **35** (9), 306–318 (Sept. 1956).
13. Hughes, T. R., et al., "To Save Energy When Reforming," *Hydrocarbon Processing*, **55** (5), 75–80 (May 1976).
14. Corneil, H. G., and Heinzelmann, F. J., "Hydrogen for Future Refining," *Hydrocarbon Processing*, **59** (8), 85–90 (Aug. 1980).
15. "Refining Handbook 1990," *Hydrocarbon Processing*, **69** (11), 83–146 (Nov. 1990).
16. Kellett, T. F., et al., "How to Select Hydrotreating Catalyst," *Hydrocarbon Processing*, **59** (5), 139–142 (May 1980).
17. Jenkins, J. H., and Stephens, J. W., "Kinetics of Cat Reforming," *Hydrocarbon Processing*, **59** (11), 163–167 (Nov. 1980).
18. McDonald, G. W. G., "To Judge Reformer Performance," *Hydrocarbon Processing*, **56** (6), 147–150 (June 1977).

19. Hartzell, F. D., and Chester, A. W., "FCCU Gets a Catalyst Promoter," *Hydrocarbon Processing*, **58** (7), 137–140 (July 1979).
20. Dale, G. H., and McKay, D. L., "Passivate Metals in FCC Feeds," *Hydrocarbon Processing*, **56** (9), 97–102 (Sept. 1977).
21. Magee, J. S., et al., "A Look at FCC Catalyst Advances," *Hydrocarbon Processing*, **58**, (9), 123–130 (Sept. 1979).
22. Howell, R. C., and Kerr, R. C., "Moving Coke? What to Expect," *Hydrocarbon Processing*, **60** (3), 107–111 (Mar. 1981).
23. Sullivan, R. F., and Meyer, J. A., "Catalyst Effects on Yields and Product Properties in Hydrocracking," presented at American Chemical Society, Philadelphia meeting, Apr. 6–11, 1975.
24. Benedek, W. J., and Mauleon, J.-L., "How First Dimersol Is Working," *Hydrocarbon Processing*, **59** (5), 143–149 (May 1980).
25. Heck, R. M., et al., "Better Use of Butenes for High-Octane Gasoline," *Hydrocarbon Processing*, **59** (4), 185–191 (Apr. 1980).
26. Logan, R. S., and Banks, R. L., "Disproportionate Propylene To Make More and Better Alkylate," *Hydrocarbon Processing*, **47** (6), 135–138 (June 1968).
27. Piel, W. J., and Thomas, R. X., "Oxygenates For Reformulated Gasoline," *Hydrocarbon Processing*, **69** (7), 68–73 (July 1990).

16

Natural Gas

R. N. Maddox* and R. E. Cannon**

CHARACTERISTICS

Natural gas is a naturally occurring mixture of simple hydrocarbons and nonhydrocarbons that exists as a gas at ordinary pressures and temperatures. In the raw state, as produced from the earth, natural gas consists principally of methane (CH_4) and ethane (C_2H_6), with fractional amounts of propane (C_3H_8), butane (C_4H_{10}), and other hydrocarbons, pentane (C_5H_{12}) and heavier.

All of the ethane and heavier components of natural gas are easily separated from the gas stream and liquefied under moderate pressure, and most of these components are recovered for their added value as separate products. Raw natural gas may be described as "rich" ("wet") or "lean" ("dry"), depending on the amounts of liquefiable heavier components contained.

In addition to hydrocarbon components, raw natural gas from some reservoirs may contain varying amounts of nonhydrocarbon contaminants or diluents, such as hydrogen sulfide (H_2S), carbon dioxide (CO_2), water (H_2O), nitrogen (N_2), and helium (He). The first two of these compounds, hydrogen sulfide and carbon dioxide, are termed "sour" or "acid" gases; acid-free natural gas is described as "sweet."

Although some natural gas is suitable for use as a fuel as produced, virtually all undergoes some processing to produce a merchantable gas whose composition is principally methane and ethane. Water and acid gas components are removed to prevent freezing, corrosion, or other operating problems in transmission or utilization. Diluents that exist in significant quantities usually are removed to improve the heating properties of the gas. Much of the ethane and propane and all of the butanes and heavier hydrocarbons are removed for their greater value as separate liquid products.

Table 16.1 is a tabulation of typical raw gas compositions; Table 16.2 shows the physical properties of the principal constituents of natural gas; and Table 16.3 depicts typical specifications for pipeline quality, or "merchantable," natural gas.

*CEC Inc., Stillwater, OK, and Emeritus Professor, Oklahoma State University.
**Executive Director, Gas Processors Association, Tulsa, OK.

TABLE 16.1 Typical Raw Gas Compositions (Water-Free Basis)

	Casinghead (Wet) Gas		*Gas Well (Dry) Gas*		*Condensate Well Gas*	
	Mol %	*Gal/mcf*	*Mol %*	*Gal/mcf*	*Mol %*	*Gal/mcf*
Carbon dioxide	0.63					
Nitrogen	3.73		1.25		0.53	
Hydrogen sulfide	0.57					
Methane	64.48		91.01		94.87	
Ethane	11.98		4.88		2.89	
Propane	8.75	2.399	1.69	0.463	0.92	0.252
iso-Butane	0.93	0.303	0.14	0.046	0.31	0.101
n-Butane	2.91	0.914	0.52	0.163	0.22	0.069
iso-Pentane	0.54		0.09		0.09	
n-Pentane	0.80		0.18		0.06	
		0.777		0.203		0.103
Hexanes	0.37		0.13		0.05	
Heptanes plus	0.31		0.11		0.06	
	100.00	4.393	100.00	0.875	100.00	0.525

Source: Gas Processors Association.

The common unit of measurement for natural gas is the cubic foot in the English system and the cubic meter in the metric system, each expressed at pressures and temperatures standard to the system. In the United States, the cubic foot is expressed in the English system at standard conditions of 14.73 pounds per square inch (psi) and 60 degrees Fahrenheit (°F), although a number of other standard conditions are specified by various regulatory and state agencies. In metric units, the cubic meter is commonly defined at one atmosphere (101.325 kPa) and 15 degrees Celsius (°C).

Another common unit of measurement for natural gas is its heating value, expressed in British thermal units (Btu) per cubic foot in the English system, and in Joules or calories in the metric system. Commercially used natural

TABLE 16.2 Properties of Natural Gas and Natural Gas Liquids

	Chemical Formula	*Molecular Mass*	*Boiling Point @14.496 psi, °F*	*Vapor Press. @100°F, psia*	*Gas Density Air = 1*	*Heating Value Btu/cf Gas*
Methane	CH_4	16.043	−258.73	(5000)*	0.554	1010.0
Ethane	C_2H_6	30.070	−127.49	(800)*	1.038	1769.6
Propane	C_3H_8	44.097	−43.75	188.64	1.523	2516.1
i-Butane	C_4H_{10}	58.123	10.78	72.58	2.007	3251.9
n-Butane	C_4H_{10}	58.123	31.08	51.71	2.007	3262.3
i-Pentane	C_5H_{12}	72.150	82.12	20.45	2.491	4000.9
n-Pentane	C_5H_{12}	72.150	96.92	15.57	2.491	4008.9
Hexane	C_6H_{14}	86.177	155.72	4.96	2.976	4755.9
Heptane	C_7H_{16}	100.204	209.16	1.62	3.460	5502.5
Octane	C_8H_{18}	114.231	258.21	0.54	3.944	6248.9
Carbon dioxide	CO_2	44.01	−109.26		1.520	0
Hydrogen sulfide	H_2S	34.08	−76.50	394.59	1.177	637.1

Source: Gas Processors Association Standard 2145.

TABLE 16.3 Typical Pipeline Quality Natural Gas

	Minimum	*Maximum*
Major and Minor Components, Mol %		
Methane	75	—
Ethane	—	10
Propane	—	5
Butanes	—	2
Pentanes and heavier	—	0.5
Nitrogen and other inerts	—	18
Carbon dioxide	—	3
Hydrogen	—	5
Total unsaturated hydrocarbons	—	0.5
Carbon monoxide	—	0.1
Trace Components		
Hydrogen sulfide	—	0.25 gr/100 scf
Mercaptan sulfur	—	0.50 gr/100 scf
Total sulfur	—	1.0 gr/100 scf
Water vapor	—	7.0 lb/mmcf
Oxygen	—	10 ppmv
Other Characteristics		
Heating value, Btu/scf—gross saturated	950	1150
—gross dry	967	1170
Relative density	0.530	0.740

Liquids: Free of liquid water and hydrocarbons at delivery temperature and pressure.

Solids: Free of particulates in amounts deleterious to transmission and utilization equipment.

gas, after processing, yields about 950 to 1050 Btu/cubic foot. Also used as a unit of measurement is the therm, equivalent to 100,000 Btu, or the nominal heat content of 100 standard cubic feet of natural gas.

The most common unit of measurement in the English system is the MCF, or thousand cubic feet. Larger volumes, used to express production or pipeline volumes, are noted as BCF, or billion (10^9) cubic feet; even larger volumes, such as reserve figures, usually are expressed as TCF, or trillion (10^{12}) cubic feet.

OCCURRENCE OF NATURAL GAS

Natural gas occurs in three principal forms: associated gas, nonassociated gas, and gas condensate. The first two of these are depicted in Fig. 16.1, which includes a simplified schematic of the flow of natural gas from the well through various stages of processing.

Associated gas is found in crude oil reservoirs, either dissolved in the crude oil or in conjunction with crude oil deposits. Dissolved or associated gas provides the driving force or reservoir pressure necessary to produce oil from a reservoir. Dissolved gases are produced from oil wells along with the oil and separates, or is separated, from the crude oil at the well head. Such gas also may be called "casinghead gas" or "oil well gas." In the beginning of the industry, virtually all utilized natural gas was associated gas from oil wells. Much of the very large gas reserves of some Persian Gulf countries and the Prudhoe Bay reserves of Alaska is associated gas.

Nonassociated gas occurs in reservoirs separate from those of crude oil. It is commonly called "gas well gas," and contains much less of the heavier, or condensable, hydrocarbons found in associated gas. Today, because of the wide demand for natural gas, most utilized gas is nonassociated gas.

Fig. 16.1. Schematic flow of natural gas from well through processing.

The third form in which natural gas reserves occur might be called a two-phase fluid, as it is neither a true gas nor a true liquid. It is not a gas because no surface boundary exists between gas and liquid. These reservoirs, called "gas condensate" reservoirs, usually are found in moderately deep formations, have very high pressures, and pose special problems in production and processing.

EVOLUTION OF THE U.S. NATURAL GAS INDUSTRY

Although natural gas is the predominant product (98%) of the gas industry today, its utilization evolved out of the production and use of manufactured gas, and its major expansion came about through utilization of early "town gas" systems originally installed to distribute manufactured gas.

Manufactured gases are those obtained by destructive distillation of coal, by the thermal decomposition of oil, or by the reaction of steam passing through a bed of heated coal or coke. Prior to the development of large natural gas reserves, manufactured gas was used widely for illumination in the "gaslight" era of the United States.

Following the discovery of major oil and gas reserves in the early twentieth century, associated gas became plentiful in oil-producing areas, and natural gas began to displace manufactured gas in town systems. However, early natural gas utilization was confined largely to areas at or very near the source of supply because methods and materials for transporting the gas had not yet been developed.

Natural gas is relatively difficult to transport and store and is extremely bulky compared to other forms of energy. At atmospheric pressure, one cubic foot of natural gas contains about two-tenths the energy content of a cubic foot of solid or liquid fuel. Thus, it is necessary to transport natural gas through pipelines at very high pressures, typically 900 to 1000 psi. The development of high-strength, thin wall pipe and modern welding techniques in the mid-1930s permitted the construction

of numerous long-distance, high-pressure gas transmission lines from the large and growing reserves of the southwestern United States to the major industrial and population centers of the country.

In rapid succession, a number of long-distance pipelines were laid during the period from 1930 to 1940, including three pioneer 1000-mile lines from large reserves in the Southwest to major markets in the North and the Northeast. Spurred by abundant and inexpensive gas reserves, the U.S. natural gas industry expanded rapidly following 1945, and natural gas virtually replaced the more expensive manufactured gas in city distribution systems.

In 1930, marketed production of natural gas totaled 1.9 trillion cubic feet; following early pipeline expansion, total consumption by 1940 had risen to 2.7 trillion cubic feet. Then, in the decade between 1940 and 1950, gas consumption more than doubled to 6.3 trillion cubic feet, and it doubled again to 12.8 trillion cubic feet by 1960. Marketed production in the United States peaked at 22.7 trillion cubic feet in 1973, followed by significant declines resulting from conservation efforts. The current marketed production of natural gas is about 18 trillion cubic feet per year, and is increasing steadily because of its convenience and qualities as an environmentally desirable fuel.

Fig. 16.2. U.S. petroleum energy production by year. (*Data from U.S. Department of Energy.*)[1,2]

Fig. 16.3. U.S. marketed natural gas production by year. (*Data from U.S. Department of Energy.*)[1]

U.S. MARKETED PRODUCTION

Natural gas accounts for well over one-half of total U.S. petroleum energy production, as illustrated in Fig. 16.2. Its major market segments, shown in Fig. 16.3, are residential and commercial fuels, industrial fuel and feedstock, and electric power generation.

Residential consumption is predominantly for space heating, cooking, and water heating. Commercial use includes space heating and other fuel needs for commercial establishments such as stores, hotels, and restaurants.

Industrial use, accounting for about 36 percent of total gas consumption, includes fuels for manufacturing operations, industrial boiler fuel, and process heat. Oil refineries are major consumers of natural gas for refining operations. In addition, natural gas is the prime feedstock for the manufacture of ammonia fertilizers, methanol, and other basic chemical derivatives. Figure 16.4 is a simplified schematic of some of the principal chemical and products derived from natural gas.

Fuel for electric power generation accounts for about 15 percent of marketed natural gas production. This market segment is extremely

Fig. 16.4. Some petrochemical products derived from natural gas. (*From Crockett, L. O., and Widgery, R., "Natural Gasoline Producers' Stake in Chemical Manufacturing," 34th Annual Convention, Gas Processors Assoc., 1955.*)[7]

price-sensitive and competes directly with fuel oil and coal. Given the current environmental emphasis on reduction of air pollutants, the use of cleaner-burning natural gas is expected to increase significantly in the foreseeable future.

In addition to what is termed "marketed production" for these uses, approximately 1.6 trillion cubic feet of natural gas is consumed annually in lease, processing plant, and pipeline fuels.

At current production levels, about 56 percent of U.S. petroleum energy production consists of the gaseous fuels natural gas and natural gas liquids. Moreover, since peak production in 1985, U.S. crude oil has declined by about 18 percent, while gaseous fuel production has increased about 6 percent. These trends suggest that the U.S. petroleum economy will become even more dependent on the gaseous fuels in the future.

NATURAL GAS LIQUIDS

Included in the gaseous fuels segment of U.S. petroleum energy production depicted in Fig. 16.2 are the natural gas liquids extracted from natural gas as produced. These natural gas liquids, commonly terms "NGLs" are ethane, propane, butanes and pentanes, and higher hydrocarbons.

The recovery and separation of natural gas liquids from raw natural gas is a relatively simple operation, which relies principally on the differences in boiling points and vapor pressures of the various components. (These

differences are shown in Table 16.2.) The two most prevalent recovery processes are the absorption and cryogenic turboexpander processes. Together, they account for approximately 90 percent of total U.S. natural gas liquids production. The bulk of U.S. gas liquids is produced as a raw mix stream in some 700 gas processing plants located in or near major gas-producing provinces. The raw mix streams are transported by an extensive pipeline network to large central fractionator complexes where the mixed stream is fractionated into its components ethane, propane, butanes, and pentanes.

The total U.S. recovery of natural gas liquids in 1990 was approximately 566 million barrels, or about 31 barrels per million cubic feet of marketed natural gas. The total NGL supply is supplemented by about 200 million barrels per year of refinery production, which equates to about 4 percent of the total crude oil charge to refineries. Thus, the total U.S. production of NGLs in 1990 was about 766 million barrels, or some 2 million barrels per day. As a point of reference, this was about 20 percent of total U.S. liquid petroleum production in 1990. The total NGL supply in 1990, including some 70 million barrels of imports, mostly from Canada, was about 836 million barrels, or 2.30 million barrels per day.

The trend of total U.S. NGL supply is depicted in Fig. 16.5. Major demand segments are shown in Fig. 16.6, which indicates that around to 35 to 40 percent of gas liquids, principally ethane and propane, are consumed as cracking feedstocks for ethylene manufacture. Some of the chemical derivatives obtained from ethane and propane are shown in Fig. 16.7.

Traditionally, a major demand for natural gas liquids, mainly butanes and pentanes, has been in the manufacture of motor gasolines. However, the recent phaseout of leaded gasoline and the current, ongoing reduction in motor gasoline volatility have drastically reduced demand for pentane and butanes as gasoline blending stocks. On the other hand, these same measures, enacted to reduce air pollutants from the automobile, have resulted in an increased demand for these products as prime feedstocks for the formulation of clean-burning reformulated gasolines.

Fig. 16.5. U.S. NGL supply by year. (*U.S. Department of Energy.*)[2]

Fig. 16.6. U.S. NGL demand by year. (*"Sales of Natural Gas Liquids and Liquefied Refinery Gases," American Petroleum Institute, Gas Processors Association, and National Propane Gas Association.*)[6]

A third major market is consumption of propane as residential and commercial fuels, principally in rural areas beyond the reach of natural gas distribution systems.

Additional major uses indicated in the "other" category of Fig. 16.6 include consumption as internal combustion engine fuels and numerous agricultural uses such as crop drying.

Fig. 16.7. Some products derived from ethane and propane. (*Crockett, L. O., and Widgery, R., Natural Gas Producers' Stake in Chemical Manufacturing," 34th Annual Convention, Gas Processors Association, 1955.*)[7]

U.S. NATURAL GAS RESERVES

Figure 16.8 is a graphical representation of proven U.S. natural gas reserves over time, which indicates approximately 10 years of indigenous supply at current consumption rates. Proven reserves estimates are subject to constant change, either up or down, depending on reservoir operating experience, field extensions resulting from additional drilling, and revised estimates. For example, an unusually large reduction in proven reserves is shown for 1988; this resulted from downward revisions in estimated gas reserves on the north slope of Alaska.

Not included in reserves data are undiscovered but probably recoverable reserves in unexplored or unproven provinces. Currently, these probable reserves are estimated by the U.S. Department of Energy at 307 to 500 trillion cubic feet, much of which almost certainly will become proven reserves when they become economically feasible.

Whatever the ultimate recoverable reserves may be, current estimates of proven reserves of natural gas comprise about 52 percent of U.S. proven reserves of total petroleum energy.

STRUCTURE OF THE U.S. NATURAL GAS INDUSTRY

The natural gas industry in the United States is made up of four major segments: production, processing, transmission, and distribution. The industry is physically interconnected by a pipeline network that extends through the lower 48 states and across international borders into Mexico and Canada.

Fig. 16.8. Proven U.S. natural gas reserves.[5]

The production function is a large part of the oil-producing industry, and most of the same firms are dominant in both oil and gas. Exploration and drilling technologies for both oil and gas are identical. Production from both oil well gas and gas well gas is introduced into large field gathering systems for delivery to the processing function. A typical gathering system may include 2000 miles of gathering lines connected to a thousand producing wells.

The gas processing function includes the gathering and delivery of gas streams from a field or fields into a central processing facility. Approximately 750 gas processing plants are in operation in the United States. In these plants the raw natural gas is dehydrated to remove moisture, treated for removal of contaminants, and compressed to pipeline pressure. During treating, most raw gas also is processed for recovery of liquid products, including ethane, propane, butanes, and natural gasoline, which are sold for petrochemical and gasoline feedstocks and other fuel uses.

The transmission function transports the processed pipeline quality gas from the gas plant to major markets. About two-thirds of U.S. marketed gas production reaches the ultimate consumer through a network of some 250,000 miles of pipelines. Other deliveries may be made directly to industrial consumers or to city distribution systems.

The distribution function receives natural gas from the transmission pipeline at the city gate of the local distribution system and delivers it to the ultimate consumer through a network totaling some 600,000 miles of distribution pipelines.

WORLD NATURAL GAS

Natural gas, in recent years, has become an international commodity fuel. Current world reserves of natural gas, depicted in Fig. 16.9,

Fig. 16.9. Estimated world natural gas reserves. (*U.S. Department of Energy.*)[4]

are estimated at about 4000 trillion cubic feet, the equivalent of about 42 percent of world petroleum energy reserves. About 38 percent of these reserves are in remote regions of the Former Soviet Union and about 31 percent are located in the Persian Gulf countries of the Middle East.

The location of these reserves, far from the major populations of consuming countries, suggests that the world natural gas industry is in a development stage very similar to that of the U.S. industry in the mid-1930s. Approximately 1.5 trillion cubic feet per year of gas from the Gröeningen field of the Netherlands, together with about 1 trillion cubic feet per year from Norway and the North Sea, is flowing by pipeline into industrialized western Europe. In addition, nearly 2 trillion cubic feet per year of natural gas is piped from the Former Soviet Union into western Europe, along with another 1.5 trillion cubic feet into eastern Europe.

Additional international movements of natural gas are in the form of liquefied natural gas (LNG), principally from Algeria into Europe (about 1 trillion cubic feet per year) and from Oceania into Japan (about 1.6 trillion cubic feet per year).

It is a certainty that world use of natural gas will increase dramatically in the immediate future as industrialized countries replace coal-fired facilities with cleaner-burning natural gas. In addition, a number of countries remote from major markets are in the process of installing world-scale plants for utilization of natural gas and gas liquids for production of fertilizers, methanol, premium gasoline blending stocks, and other basic petrochemical derivatives that will result in higher-value products from natural gas and gas liquids for which these are no local markets.

Already, the gaseous fuels account for approximately 40 percent of world petroleum energy production (Fig. 16.10) and nearly

Fig. 16.10. World petroleum energy production. (*U.S. Department of Energy.*)[4]

one-half of world petroleum energy reserves. And, as illustrated in Fig. 16.11, the growth of world natural gas consumption is already increasing rapidly.

Despite a world surplus of crude oil capacity for the foreseeable future, it seems certain that the world has entered the "Age of Gaseous Energy," and that natural gas and gas liquids will continue to move into a position of petroleum energy dominance in the next decade.

PREPARING NATURAL GAS FOR TRANSMISSION AND SALE

To make a product suitable for sale, the various contaminants and undesirable constituents contained in the raw gas must be removed. The first step is to separate the gas from any liquids and/or solids. This initial separation is usually carried out in near proximity to the producing well. The separators used depend primarily upon the difference in density between the gas and the liquid for separating the two phases. These separators may be classified as vertical, horizontal, or spherical, depending on their geometry and physical orientation.[8]

Fig. 16.11. World gas consumption. (*U.S. Department of Energy.*)[4]

The liquids produced are sold as oil or condensate, depending upon the type of reservoir from which the production is obtained. The natural gas stream after this first stage of separation contains only gaseous components, though some of these may be liquefied by later processing.

Typical compositions for gas streams at this point are shown in Table 16.1. (The compositions in Table 16.1 are on a water-free basis.) As typically produced, each of the gas streams would be saturated with water vapor. Developing the technology for handling gas streams that contain significant quantities of water vapour was one of the keys necessary to development of the natural gas industry as it exists today.

In the presence of water, natural gas can—under the proper conditions of temperature and pressure—form hydrates. To the naked eye the hydrates are similar in appearance to snow; and they can result in plugging of pipelines and flow systems. Hydrates can form at temperatures well above the freezing point of water. In the early days of pipelining natural gas under pressure, shutdowns of the lines during the winter months frequently were caused by plugging with hydrates.

A hydrate is a type of chemical compound called a clathrate, defined as a type of solid molecular compound in which one component is trapped in cavities of cagelike crystals of another component. In the natural gas hydrate, water forms the "cage," and hydrocarbon molecules are the trapped component.

Hydrate formation can be prevented or avoided in two ways: (1) by dehydration, in which water is removed from the gas stream so that under conditions of processing or transmitting, no liquid water will condense; and (2) by inhibition, which involves injecting into the gas stream a component that will interfere with the ability of the water to form hydrates with the gas. Both procedures are widely used in the gas industry.

Of the wide variety of materials that can effectively depress the hydrate-forming tendency of a natural gas, two—methyl alcohol and ethylene or diethylene glycol—are used almost exclusively. Ethylene glycol (a polyhydroxy alcohol) is a high-boiling point material that mixes in all proportions with water and remains predominantly in the liquid phase. At temperatures approaching the freezing point of water, solutions containing more than 50 percent by weight ethylene glycol have viscosities approaching 10 cp. This limits their utilization in some low-temperature applications.

Methyl alcohol has a much lower boiling point and consequently a much higher vapor pressure than ethylene glycol. Hence procedures used for estimating the amount of methyl alcohol required must take into account the amount of methyl alcohol that vaporizes into the gas phase.[9,13,14] Although the vaporization requires the injection of greater quantities of methanol, it does have potential advantages. If the gas phase were cooled in further processing or transmission, one would expect methanol to condense along with additional water and/or hydrocarbons. This would provide a safety factor for hydrate prevention that does not exist with the glycols.

Dehydration of the gas can be accomplished through the use of either a liquid or a solid desiccant. The solid desiccants most often employed include alumina, silica gel, and molecular sieves. The liquid desiccant most frequently used is triethylene glycol.

Any of the solid materials used as desiccants must be regenerated periodically; so there must be at least two adsorption towers if dehydration of the gas is to be continuous. A schematic of a typical two-tower adsorption plant is shown in Fig. 16.12.[9] The adsorbent must be heated to high temperature (ca. 400°F, 200°C) to remove the water and regenerate the solid desiccant. This usually is accomplished by heating a slip stream of the gas and circulating the heated gas through the adsorbent bed.

Figure 16.13 shows a schematic flow sheet for a typical triethylene glycol dehydration system. The lean glycol is pumped to the top of an absorber column and flows downward in countercurrent contact with the water-wet gas entering the bottom of the tower. The absorber may contain either plates or packing. Dehydrated gas leaves the top of the absorber,

Fig. 16.12. Two-tower adsorption dehydration unit.

and the glycol enriched by the absorbed water flows from the bottom into a regenerator. After the absorbed water is stripped from the triethylene glycol solution, it is recirculated to the absorber. The concentration of the glycol leaving the reboiler will be approximately 98.5 to 98.7 percent by weight triethylene glycol if the reboiler operates at 400°F (205°C) and one atmosphere pressure.

Normal contract specfication for gas to be transmitted through a high pressure pipeline is a water content of 7 lb of water per million standard cubic feet of natural gas. This is approximately the water content of natural gas in equilibrium with water at the freezing point when the gas is under a pressure of 1000 psia.

Fig. 16.13. Glycol dehydration unit.

The solid desiccants offer much lower water contents for the dried gas. With proper design and operation, molecular sieve dehydrators can satisfactorily prepare gas for total liquefaction at temperatures as low as −263°F (−165°C). Lower outlet water dew points can be obtained with triethylene glycol if the concentration of glycol going to the absorber is increased. This can be accomplished by using vacuum in the reboiler or by introducing another material into the reboiler to lower the required partial pressure of the water vapor above the glycol. Proper design and operation of such glycol systems can produce outlet water dew points in the −40°F (−40°C) range, or lower.

Natural gas pipelines normally operate at elevated pressures. In many instances the gas is available at low pressures so must be pumped or compressed to higher pressure. If the volume of the gas stream is small, reciprocating compressors must be used. For larger

Fig. 16.14. Approximate range of application for compressors. (GPSA Engineering Data Book, *10th ed., Gas Processors Suppliers Association, Tulsa, OK, 1987.*)

gas volumes, rotary compressors of various types are available. Figure 16.14[10] shows pressure and flow rate ranges in which different types of compressors currently available are used.

If the natural gas stream contains objectionable quantities of hydrogen sulfide and/or carbon dioxide, they must be removed in order to make the gas suitable for transmission and sale. The details of removal of H_2S and CO_2 from natural gas streams are beyond treatment here, but excellent discussions are available.[10,12] There are many different processes available, depending upon the contaminants to be removed and their concentration in both the sour and the sweetened gas. The dominant treating process is still the use of an alkanolamine. A typical flow sheet for an amine sweetening installation for removal of hydrogen sulfide in carbon dioxide from a natural gas stream is shown in Fig. 16.15.[12]

Environmental restrictions and controls will not permit release of the hydrogen sulfide or its incineration to sulfur dioxide. As pointed out by Maddox,[12] the least undesirable alternative is conversion of the hydrogen sulfide to sulfur. Again, there are many different processes available, several of which are discussed in the book by Maddox. The dominant method is by use of the Claus process, which involves partial combustion of the hydrogen sulfide to sulfur dioxide and catalytic combination of the hydrogen sulfide and sulfur dioxide to produce sulfur. Detailed discussions of the Claus technology as well as other techniques of converting hydrogen sulfide to sulfur are available.[10,12]

Figure 16.16 is a sketch of a "once-through" Claus process using a hot gas bypass for reheat. Sufficient air is used to burn one-third of the H_2S and all hydrocarbons. Steam usually is generated in the waste heat boiler, where the combustion products are cooled, and sulfur is condensed. The gas must be heated before introduction to the reactor to prevent sulfur condensation in the reactor

Fig. 16.15. Typical ethanolamine sweetening unit.

Fig. 16.16. Once-through Claus sulfur unit.

bed. Tail gas usually is sent to a tail gas cleanup unit to minimize sulfur emissions.

PROCESSING FOR LIQUIDS RECOVERY

The first time that liquids were recovered from a natural gas stream, it probably happened more by accident than by design. Any time a gas stream from a liquid-gas separator is either cooled or compressed, liquids will form. This simple statement still provides the basis for nearly all hydrocarbon liquid recovery systems from natural gas. Figure 16.17[11] shows schematically the flow through a simple compression liquids-recovery facility. Replacing the water cooling after the second stage of compression with refrigeration would increase liquids recovery. If the gas had not been dehydrated, there would be danger of hydrate formation at this point.

As the natural gas industry developed and expanded, so did the demand for liquids recovered from natural gas. One way of recovering additional liquids is through the use of a heavy absorption oil that has good solubility characteristics for the propane and heavier hydrocarbon components in the gas stream. As in the case of dehydration, the solvent must be regenerated. Figure 16.18[11] shows the absorption–stripping section of an absorption-type hydrocarbon liquids recovery facility. The absorbent is circulated to the top of the absorber and flows downward, absorbing liquefiable components from the gas stream. It then is heated in the stripper, with the stripped components being cooled and condensed while the stripped absorption oil is recirculated. The liquid hydrocarbon stream typically would flow through a fractionation train for separation into the individual components or mixtures that could be sold.

Fig. 16.17. Liquid hydrocarbon recovery by compression and cooling. (*Adapted from Huntington, R. L.,* Natural Gas and Gasoline, *McGraw-Hill, New York, 1950.*)

Fig. 16.18. Absorption-stripping unit. (*Adapted from Huntington, R. L.,* Natural Gas and Gasoline, *McGraw-Hill, New York, 1950.*)

Combining refrigeration with oil absorption allowed for additional liquid recoveries and greater economy of operation. Using temperatures as low as $-40°F$ ($-40°C$), refrigerated oil absorption plants were capable of recovering 70 percent or more of the propane present in the natural gas stream. Recoveries for the butane and heavier constituents were essentially 100 percent.

Development of the turbo-expander process allowed design and construction of plants for recovery of liquid ethane, as well as the heavier hydrocarbon components. The turbo-expander extracts work from the gas during expansion from a high pressure to a lower pressure. Because of the work extraction the gas is cooled, and, by means of suitable heat exchange, temperatures as low as $-150°F$ ($-100°C$) can easily be achieved. The separation at low pressure gives higher relative values of the vapor–liquid equilibrium constant for nitrogen and methane than for ethane and heavier hydrocarbons. The result is that the process shown in Fig. 16.19[13] can recover as much as 75 percent or more of the ethane contained in the gas as the demethanized product. In some cases the process will be run

Fig. 16.19. Turbo-expander process flow.

to produce a deethanized product that contains essentially all of the propane in the feed gas. The liquid product from the facility either can go to on-site fractionation into salable products or, more typically, be introduced into a pipeline and transported to a central fractionation facility near the point of ultimate use for the liquid product.

REFERENCES

1. *Natural Gas Monthly*, Energy Information Administration, U.S. Department of Energy.
2. *Petroleum Supply Monthly*, Energy Information Administration, U.S. Department of Energy.
3. *Petroleum Supply Annual*, Energy Information Administration, U.S. Department of Energy.
4. *International Energy Annual*, Energy Information Administration, U.S. Department of Energy.
5. "U.S. Crude Oil, Natural Gas and Natural Gas Liquids Reserves," Energy Information Administration, U.S. Department of Energy.
6. "Sales of Natural Gas Liquids and Liquefied Refinery Gases," American Petroleum Institute, Gas Processors Association, and National Propane Gas Association.
7. Crockett, L. O., and Widgery, R. C., "Natural Gasoline Producers' Stake in Chemical Manufacturing," *Proc. 34th Annual Convention, Gas Processors Association*, 1955.
8. Campbell, J. M., *Gas Conditioning ad Processing Vol. 1—The Basic Principles*, 6th ed., Campbell Petroleum Series, Norman, UK, 1984.
9. Campbell, J. M., *Gas Conditioning and Processing Vol. 2—The Equipment Modules*, 6th ed., Campbell Petroleum Series, Norman, OK, 1984.
10. *GPSA Engineering Data Book*, 10th ed., Gas Processors Suppliers Association, Tulsa, OK, 1987.
11. Huntington, R. L., *Natural Gas and Natural Gasoline*, McGraw-Hill, New York, 1950.
12. Maddox, R. N., *Gas Conditioning and Processing Vol. 4—Gas and Liquid Sweetening*, 3rd ed., Campbell Petroleum Series, Norman, OK, 1985.
13. Maddox, R. N., and Lilly, L. L., *Gas Conditioning and Processing Vol. 3—Computer Applications for Production/Processing Facilities*, 2nd ed., Campbell Petroleum Series, Norman, OK, 1990.
14. Segeler, C. G., (Editor-in-Chief), *Gas Engineers Handbook*, 5th ed., 5th printing, Industrial Press, Inc., New York, 1977.

17

Coal Technology

M. Rashid Khan,[1] B. K. Parekh,[2] Michael A. Serio,[3] G. Todd Hager and Frank J. Derbyshire,[4] and Victor H. Agreda[5]

INTRODUCTION

Coal represents over 90 percent of the U.S proven reserves of fossil fuels, and can serve as a source of synthetic fuel for the petrochemical industry as well as a source of electric power production and process heat generation. About one-third of the world's coal reserves are present in the United States. The recoverable reserves of U.S. coal have been estimated to be 250 billion tons. Eastern U.S. coals are generally bituminous with a heating value of 10,000 to 15,000 Btu/lb. Bituminous coals comprise nearly one-half of the total U.S. coal reserves. The western as well as southwestern U.S. coals are mainly subbituminous (with a heating value of 9000–12,000 Btu/lb) and lignite (with a heating value of 8000–10,000 Btu/lb).

Historical interest in coal technology has been cyclic, primarily because of the varying economics of petroleum. Coal dominated the U.S. energy picture until the discovery of petroleum. Each subsequent threat to oil supplies spurred a temporary renewal of interest in coal, which rapidly faded as each new major oil strike was made. The threat to oil supplies in 1973 produced the historic response pattern in all respects except one: although exploration for oil and gas increased because of the price increases, few major new oil reserves have been found. Coal use for electric power generation has increased in recent years, but the economic conversion of coal to clean fuels still stands as one of the key technical challenges of our time.

The number of publications on the subject of coal composition, structure, and usage has grown enormously during the last few years, as funds for coal research have increased markedly. Indeed, the future of coal science looks brighter than ever as a new generation of researchers grapple with, and overcome the obstacles to, making meaningful contributions toward the elucidation of the structure, composition, and physiocochemical behavior of coal. The objective of this chapter is to provide a brief review of the state of the art of coal science and technology.

(1) Texaco Research and Development: Coordinating author plus Introduction and sections on Coal Production and Consumption, Origin and Classification, Coal Composition and Structure, Pyrolysis, and Gasification; (2) Center for Applied Energy Research, University of Kentucky: Coal Mining and Preparation; (3) Advanced Fuel Research: Combustion; (4) Center for Applied Energy Research, University of Kentucky: Direct Coal Liquefaction; (5) Eastman Chemical Company: Petrochemical Feedstocks.

Fig. 17.1. U.S. coal production, 1949–90. (Annual Energy Review, *Energy Information Administration, U.S. Department of Energy, DOE/EIA-0384 (90).*)[1]

COAL PRODUCTION AND CONSUMPTION

Bituminous coal accounts for, by far, the largest share of all coal produced in the United States. In 1990, production of all types of coal totaled over 1 billion short tons, of which 943 million tons were subbituminous and bituminous coals (Fig. 17.1). Despite its clean burning characteristics, anthracite accounts for a diminishing share of total coal production. Most anthracite is mined in northeastern Pennsylvania. In 1949, anthracite accounted for 9 percent of total coal production; by 1990 its share was reduced to less than one-half of 1 percent. Figure 17.2 presents data on coal consumption in the United States.

Although more coal is mined to the east of the Mississippi than in the West, the share of western production has increased significantly since 1965. In 1965, production of western coal was 27 million tons (5% of the total). By 1990, western production had increased by almost 15 times, to 401 million tons (39% of the total). The growth in western coal, due in part to environmental concerns, led to an increased demand for low-sulfur coal, which is concentrated in the West. Also, surface mining, with its higher than average productivity, is much more prevalent in the West than the East.

The average productivity of all types of mines in the United States increased each year after 1949, reaching 2.4 tons per miner per hour in 1969. In 1989, average productivity reached an all-time high of 3.7 tons per miner hour.

Electric utilities are the dominant consumers of coal. Their consumption grew from 84 million tons in 1949 (a 17% share) to 772 million tons (an 86% share) in 1990. In contrast, consumption by all other sectors in 1990 was lower than it had been in 1949. The most dramatic declines occurred in the transportation sector (as railcars switched to petroleum) and the residential and commercial sectors. In 1949, those two sectors accounted

Fig. 17.2. U.S. coal consumption by sector. (*U.S. Department of Energy, DOE/EIA-0384 (90).*)[1]

Fig. 17.3. World coal production in 1989. (*U.S. Department of Energy, DOE/EIA-0384 (90).*)[1]

for 187 million tons (39% of the total); but in 1990, their consumption totaled 6 million tons, less than 1 percent of U.S. consumption. Coal consumption by the industrial sector trended downward after the mid-1960s. Industrial consumption fell to about 112 million short tons in 1986 from 205 tons in 1966. During 1988 and 1989, coal consumption rose to 118 tons, but it declined to 115 tons in 1990 during an economic slowdown in the United States. As shown in Fig. 17.3, the United States is the second largest producer of coal in the world, right behind China.

ORIGIN AND CLASSIFICATION OF COAL

Coal originated from the remnants of various forms of plant life (trees, bushes, ferns, mosses, vines, etc.) that flourished in swamps and bogs millions of years ago during prolonged periods of abundant rainfall. Peat, the precursor of coal, was formed by bacterial and chemical action on accumulations of plant debris. Given[2] has presented an excellent review on the origin and formation of coal.

Accumulations of plant debris and sediments compressed and solidified the deposit, thus initiating coalification. Exclusion of oxygen from the peat by overlying water retarded the rate of decay, and the extent of coalification (which determines the coal rank) depended largely on the temperatures existing during this process. When the prevailing temperature was relatively low (corresponding to a depth of only a few hundred feet below the surface), coalification was relatively slow and resulted in low-rank coal. At increased depths, higher temperatures in the organic masses improved the coalification process, producing higher-rank coals (bituminous and anthracite). Intrusions of igneous rocks sometimes accelerated these increases. The chemical changes that occur during coalification are decreases in hydrogen and oxygen content and an increase in carbon content.

Coal is an "organic rock," being grossly characterized by layers or bands having glossy or dull appearances. The glossy layers are composed mainly of vitrinite, formed from woody parts of plants: trunks, limbs, branches, twigs, and roots. The dull bands consist of finely divided material, formed from leaves, pollen, spores, seeds, and resins (collectively termed exinite), in addition to fine vitrinite and more coalified micrinite and fusinite. Besides these banded coals, there are two types of dull, nonbanded coals: (a) cannels, which are rich in spores, and (b) bogheads, which contain abundant remains of algae.

Coals can be classified by rank and/or type. Table 17.1 presents a classification of coals in the United States according to rank. The main rank classes seen in this table are grouped according to certain chemical and physical properties. The factors determining rank, or degree of coalification, are moisture, volatile matter (material that is volatilized when coal is heated at a certain temperature and for a certain length of time), fixed carbon (the residue after the loss of volatile matter), heating value, caking, and weathering properties.

The reflectance of vitrinite (a commonly used rank indicator for high-rank coals) is measured by comparing the intensity of a beam of light incident on a polished surface of vitrinite with the light reflected directly back from that surface.

An international classification of coal (Table 17.2) classifies high-rank coals according to their volatile-matter content, calculated on a dry, ash-free (d.a.f.) basis. Because volatile matter is not a suitable parameter for coals containing more than 33 percent volatile matter, the calorific value on a moist, ash-free basis is included as a parameter for such coals. The resulting classes of coal, based on volatile-matter content and calorific value, are grouped according to their plastic properties, by employing either the free-swelling or the Roga test. They are then subgrouped according to plastic properties, by using either the Audibert-Arnu or the Gray-King test. A three-figure cost number is used to classify a coal. The first figure indicates the class of the coal, the second figure the group, and the third figure the subgroup.

TABLE 17.1 Classification of Coals by Rank[a]

Class	Group	*Fixed Carbon Limits, percent (Dry, Mineral-Matter-Free Basis)* Equal or Greater Than	Less Than	*Volatile Matter Limits, percent (Dry, Mineral-Matter-Free Basis)* Greater Than	Equal or Less Than	*Calorific Value Limits, Btu per pound (Moist,[b] Mineral-Matter-Free Basis)* Equal or Greater Than	Less Than	*Agglomerating Character*
I. Anthracitic	1. Meta-anthracite	98	—	—	2	—	—	—
	2. Anthracite	92	98	2	8	—	—	—
	3. Semianthracite	86	92	8	14	—	—	Nonagglomerating[c]
II. Bituminous	1. Low volatile bituminous coal	78	86	14	22	—	—	
	2. Medium volatile bituminous coal	69	78	22	31	—	—	
	3. High volatile *A* bituminous coal	—	69	31	—	14,000[d]	—	Commonly agglomerating[e]
	4. High volatile *B* bituminous coal	—	—	—	—	13,000[d]	14,000	
	5. High volatile *C* bituminous coal	—	—	—	—	11,500	13,000	
						10,500	11,500	Agglomerating
III. Subbituminous	1. Subbituminous *A* coal	—	—	—	—	10,500	11,500	Nonagglomerating
	2. Subbituminous *B* coal	—	—	—	—	9,500	10,500	—
	3. Subbituminous *C* coal	—	—	—	—	8,300	9,500	—
IV. Lignitic	1. Lignite *A*	—	—	—	—	6,300	8,300	—
	2. Lignite *B*	—	—	—	—	—	6,300	—

[a]This classification does not include a few coals, principally nonbanded varieties, which have unusual physical and chemical properties and which come within the limits of fixed carbon or calorific value of the high-volatile bituminous and subbituminous ranks. All of these coals either contain less than 48 percent dry, mineral-matter-free fixed carbon or have more than 15,500 moist, mineral-matter-free British thermal units per pound.

[b]Moist refers to coal containing its natural inherent moisture but not including visible water on the surface of the coal.

[c]If agglomerating, classify in low-volatile group of the bituminous class.

[d]Coals having 69 percent or more fixed carbon on the dry mineral-matter-free basis shall be classified according to fixed carbon, regardless of calorific value.

[e]It is recognized that there may be nonagglomerating varieties in these groups of the bituminous class, and there are notable exceptions in high volatile *C* bituminous group.

TABLE 17.2 International Classification of Coals by Type (Courtesy Bureau of Mines, U.S. Department of the Interior)

GROUPS (determined by caking properties)			CODE NUMBERS										SUBGROUPS (determined by coking properties)		
GROUP NUMBER	ALTERNATIVE GROUP PARAMETERS		The first figure of the code number indicates the class of the coal, determined by volatile-matter content up to 33% V. M. and by calorific parameter above 33% V. M. The second figure indicates the group of coal, determined by caking properties. The third figure indicates the subgroup, determined by coking properties.										SUBGROUP NUMBER	ALTERNATIVE SUBGROUP PARAMETERS	
	Free-swelling index (crucible-swelling number)	Roga index												Dilatometer	Gray-King
3	>4	>45					435	535	635				5	>140	>G8
						334	434	534	634				4	>50-140	G5-G8
						333	433	533	633	733			3	>0-50	G1-G4
						332 a / 332 b	432	532	632	732	832		2	≤0	E-G
2	2½-4	>20-45				323	423	523	623	723	823		3	>0-50	G1-G4
						322	422	522	622	722	822		2	≤0	E-G
						321	421	521	621	721	821		1	Contraction only	B-D
1	1-2	>5-20			212	312	412	512	612	712	812		2	≤0	E-G
					211	311	411	511	611	711	811		1	Contraction only	B-D
0	0-½	0-5		100 A / 100 B	200	300	400	500	600	700	800	900	0	Nonsoftening	A
CLASS NUMBER →			0	1	2	3	4	5	6	7	8	9	As an indication, the following classes have an approximate volatile-matter content of: Class 6 33-41% volatile matter; 7 33-44% " "; 8 35-50% " "; 9 42-50% " "		
CLASS PARAMETERS: Volatile matter (dry, ash-free) →			0-3	>3-10; >3-6.5; >6.5-10	>10-14	>14-20	>20-28	>28-33	>33	>33	>33	>33			
CLASS PARAMETERS: Calorific parameter [a] →			–	–	–	–	–	–	>13,950	>12,960-13,950	>10,980-12,960	>10,260-10,980			
CLASSES (Determined by volatile matter up to 33% V. M. and by calorific parameter above 33% V. M.)															

Note: (i) Where the ash content of coal is too high to allow classification according to the present systems, it must be reduced by laboratory float-and-sink method (or any other appropriate means). The specific gravity selected for flotation should allow a maximum yield of coal with 5 to 10 percent of ash.
(ii) 332a . . . >14-16% V. M.
332b . . . >16-20% V. M.

[a] Gross calorific value on moist, ash-free basis (30°C, 96% relative humidity) B.t.u./lb.

COAL COMPOSITION, ANALYSES, AND STRUCTURE

Coal Composition and Analyses

Coal is an "organic rock" composed of carbon, hydrogen, oxygen, nitrogen, sulfur, and mineral matter. These components can vary considerably both chemically and structurally, depending on the rank of the coal and even within coals of the same rank. Of these parameters, sulfur and mineral matter exhibit the greatest variation for coals of the same rank. For coals of different ranks the greatest characteristic variations occur in the carbon and hydrogen values.

There are two major types of coal analysis—proximate and ultimate. In addition, there are a number of miscellaneous analyses. Many of the tests are empirical, requiring strict adherence to specified conditions. In the United States, the American Society for Testing and Materials (ASTM) develops these methods. International standards are developed by the International Organization for Standardization.

Proximate analysis includes the determination of moisture, ash, and volatile matter (the fixed carbon by difference). These values, in addition to total sulfur and heating values, usually are sufficient to provide the information necessary for process heating, steam generation, and coking considerations. The ultimate analysis includes the determination of carbon, hydrogen, nitrogen, and total sulfur (ith oxygen being calculated by difference). Numerous other analyses are conducted, the scope of which varies widely. These include not only standard methods such as heating value, free-swelling index, and ash analysis but also those procedures used to determine trace elements, mineral phases, surface patameters, conversion potential, and plastic properties.

The standard analyses are described in ASTM and Bureau of Mines publications. Many of the procedures, however, are not standard methods and must be acquired from one or more of the many publications dealing with the specific area of interest. For example, plastic properties of coal, which are important considerations in any process design, can be characterized by various techniques. Each technique has its own advantages and disadvantages (Table 17.3).

Figure 17.4 shows the proximate analyses of coals selected to represent the various ranks of coal (reported on an ash-free basis). Analyses listed by coal bed, county, and state can be found in a number of publications, such as the data bank developed by the U.S. Department of Energy and the Pennsylvania State University Coal Bank, Bureau of Mines bulletins, Illinois State Geological Survey circulars, and the Argonne National Laboratory Coal Data Base (on a limited number of coal samples).

The heating value (usually reported in Btu/lb) is the high-heat value (HHV), also referred to as the gross heating value. This quantity is defined as the heat produced by combustion of a unit quantity, at a constant volume, under specified conditions. The low-heat value (LHV) is calculated from the HHV by deducting 1030 Btu for each pound of water originally present or formed during combustion. The high-heating value can be reasonably calculated from a number of formulas using determined values for carbon (C), hydrogen (H), oxygen (O), sulfur (S), and additional data as deemed necessary.

Coal Structure

An obstacle to progress in coal technology is that most scientists are accustomed to working with homogeneous species of good purity, but there is no such thing as pure coal. The composition and the structure can vary dramatically between coals of different ranks, or even between coals of the same rank. However, some structural features are common to all coals.[88]

In general, coals are composed of aromatic and hydroaromatic building blocks or layers, with a variety of substituent groups and some cross-linking between adjacent units. Some of these units are not strictly planar because of the presence of heteroatoms (oxygen, nitrogen, or sulfur) or hydroaromatic portions. The structure of graphite consists of extensive parallel layers of carbon atoms arranged

TABLE 17.3 Summary of Various Techniques to Characterize the Plastic Properties of Coal

	Description	*Advantages*	*Disadvantages*
Free-Swelling Index/Grey-King	Visual comparison of cokes carbonized under well-defined conditions, with standard conditions.	Inexpensive, rapid, reproducible.	Lack of flexibility; transition phenomena from coals to coke not known.
Hot-Stage Microscopy	Morphological changes observed by optical microscopy.	Appropriate to supply the changes in macerals when heated.	Provides qualitative information "subjective."
Foxwell/Gas Flow Method	Measures the resistance of the gas flow through a bed of coal during continuous heating.	Indirectly measures the fluidity of the plastic mass by the pressure differences between the inlet and outlet ports.	Requires highly purified (oxygen-free) gas to avoid oxidation.
Plastometer a. Constant torque b. Variable torque	a. Records the angular velocity of a rotating shaft through a bed of coal as a function of temperature. b. Records the changes in torque re-required to rotate a cylindrical retort (when coal is carbonized) at constant speed as a function of temperature.	Measures the changes in fluidity as a function of temperature; quite sensitive for determining extent of oxidation or weathering.	Does not provide data on swelling properties, limited heating rates ($-3°C/min$).
Dilatometer	Expansion, contraction, and volume change on resolidification measured as a function of temperature when heated.	Provides data on dilation parameters and transition temperatures; wide range of heating rates ($<100°C/min$).	Does not provide quantitative information on changes in fluidity of coal.

Fig. 17.4. Analysis of U.S. coals selected to represent the various ranks. (*Courtesy Bureau of Mines*).

in condensed layers mutually oriented and separated by a distance of 0.335 μm (the carbon–carbon distance of van der Waal). The parallel stacking of layers in coal is similar to this, but lacks mutual orientation between the layers, and the average spacing between layers is somewhat larger in coal than in graphite.

Low-rank coals are composed of small layers that are more or less randomly oriented and are connected by cross links. Thus, the structure is highly porous. Medium-rank coking coals shows a greater degree of orientation, and a greater tendency toward parallel stacking. These coals have fewer cross links and fewer pores than lignites. High-rank coals (primarily anthracites) show a growth of the individual layers, a marked increase in the degree of orientation, and the development of a new type of pore that is elongated parallel to the stacks of layers. These coals are considered to be the products of the highest degree of coalification.

The above classification of coal types must be used with caution. In essence, the heterogeneous chemical structures of the wide range

of plant materials and chemicals that were the genesis of coal guarantee an almost unlimited range of chemical structures within the various types of coals. Therefore, coals have been considered as a variety of chemical entities; and this wide variety virtually dictates coal reactivity under specific conditions.

The identification of the constituents of a complex mixture such as coal by molecular type may proceed in a variety of ways, but it generally consists of three types of analysis: chemical techniques; spectroscopic techniques; and physical property methods, whereby structural parameters are derived from a particular property by a sequence of mathematical manipulations. Several good reviews on this subject are available.[3,4]

Relations between Elemental Composition of Coal and Environmental Concerns

Coals generate "a seemingly endless string of environmental hazards that are ubiquitous and pervasive."[5] It is well known that nitrogen and sulfur are the two primary pollutants (heteroatoms) in coal. Therefore, the structure of these compounds in coals, their evolution during thermal decomposition, and the distribution of these species in products are being investigated. The studies conducted in these areas suggest that nitrogen in coal is almost entirely contained in tightly bound ring structures. The nitrogen-containing structures evolve from coal without cleavage during pyrolysis. Coal-derived nitrogen compounds also result in the formation of materials that interfere with synthetic fuel processing. The latter problem results from the fact that most of these compounds are basic and thus serve as poisons for the acidic catalysts used in synthetic crude oil refining. The nitrogen present in coal is converted to NO_x compounds during combustion.

Sulfur is present in coal in several forms, such as pyrite, sulfates (small amounts), and organic-sulfur compounds. Pyrite (FeS_2) decomposes at a relatively low temperature (approximately 400°C), producing FeS, H_2S, and organic-sulfur compounds. Organic-sulfur compounds can be classified into three groups: loosely bound sulfur, tightly bound sulfur, and a category intermediate between these two forms. The tightly bound sulfur compounds decompose above 500°C, and loosely bound organic-sulfur decomposes at low temperatures (about 400°C). Some of the tightly bound organic-sulfur compounds are formed by incorporating pyritic sulfur into the ring structures.

Pollutant distributions between char, tar, and gases are expected to be a function of pyrolysis conditions and, hence, are process-dependent. Khan[6] developed models to predict the distribution of sulfur in the products based on pyrolysis data for 32 different coal samples devolatilized in a fixed-bed reactor. It was observed that the total sulfur content in the pyrolysis products at 500°C could be correlated with the total sulfur (weight percent of original coal) content. The relationships are as follows:

$$\text{Total sulfur in gas} = 0.31 \times S_{coal}$$
$$[R^2 = 0.93, F = 425, P = 0.0001]$$

$$\text{Total sulfur in tar} = 0.06 \times S_{coal}$$
$$[R^2 = 0.85, F = 175, P = 0.0001]$$

$$\text{Total sulfur in char} = 0.61 \times S_{coal}$$
$$[R^2 = 0.98, F = 1{,}475, P = 0.0001]$$

where R^2 is the coefficient of determination, F is the mean square of the model divided by the mean square of error, and P is the significance value. When $P < 0.05$, the effect usually is termed significant.

Multivariate analysis demonstrated that the organic sulfur content of coal plays a stronger role in determining the sulfur content of tar or gases, whereas the pyritic sulfur content plays a key role in determining the sulfur content of the char, at the low-temperature (500°C) devolatilization conditions. With increasing peak devolatilization temperature, however, the gaseous sulfur yield increases at the expense of char sulfur.[6]

COAL MINING AND PREPARATION

Coal mines fell into two general classifications, surface and underground. Production of coal

Fig. 17.5. U.S. coal production by methods of mining. (*Sources: DOE/EIA* Annual Energy Review *1989,* Annual Energy Outlook *1991, and* Coal Data.)

by type of mining for the United States is shown in Fig. 17.5.

Surface Mining

Surface mining techniques are used when the coal is present near the surface, and the overburden is thin enough. These techniques include contour mining, strip mining, and auger mining.

Contour mining is used in hilly countryside areas where the slope of the surface will permit only a narrow bench to cut around the side of a hill; the excavation is backfilled immediately after the removal of coal. It is the only method that can be used on slopes of 15 degrees or higher.

Strip mining is used in flat or gently rolling lands on the Midwest and West where large and efficient equipment can be used. In this

technique, the coal is exposed by removing the overlying strata, or overburden. Blast holes are drilled, and explosives are loaded into these holes to shatter the rock cover; earth-moving equipment is used to remove the soil and the shattered rock. The coal then is collected with power shovels or other coal-digging machines and loaded directly into trucks.

Auger mining is a supplementary method used to reach coal in stripped areas where the overburden has become too thick to be removed economically. Large augers are operated from the floor of the surface mines and bore horizontally into the coal face to produce some reserves not otherwise minable.

For a coal deposit where the seam is near the top of a hill, the entire hilltop can be removed to expose the coal.

Underground Mining

Underground mining techniques are somewhat more labor-intensive than surface mining and are used to remove coal located below too much overburden for surface mining; but here, too, machines are used in most instances to dig, load, and haul the coal. Access to the coal seam is through a drift (horizontal passage), a slope, or a shaft (Fig. 17.6), depending on the location of the coal seam.

A *drift mine* is one that enters a coal seam exposed at the surface on the side of a hill or mountain. The mine follows the coal horizontally.

A *slope mine* is one where an inclined tunnel is driven through the rock to the coal, with the mined coal removed by conveyors or truck haulage.

A *shaft mine* is one where a vertical shaft is dug through the rock to reach the coal, which may be at great depth below the surface. The coal then is mined by horizontal entries into the seam, with the resulting coal hoisted to the surface through the vertical shaft.

Two general mining systems are used in underground mining, namely, room-and-pillar and longwall.

Room-and-pillar mining is an open-stopping method where mining progresses in a nearly horizontal or low-angle direction by opening multiple stopes or rooms, leaving solid material to act as pillars to support the vertical load. This system recovers about 50 percent of the coal and leaves an area much like a checkerboard. It is used in areas where the overlying rock or "roof" has geologic characteristics that offer the possibility of good roof support. This system was used in the old mines where the coal was hand-dug. Two current methods for extracting the coal from the seam are the conventional method, where the coal is undercut and blasted free (Fig. 17.7), and the continuous method, where a machine moves along the coal face to extract the coal instead of blasting it loose (Fig. 17.8). Roof control is the major problem of the room-and-pillar method of mining.

Longwall mining uses a machine that is pulled back and forth across the face of the coal seam in larger rooms. It is not a major mining method in the United States at this time. Coal recovery using this method is greater than in room-and-pillar mining and

Fig. 17.6. Three types of entrances to underground mines—shaft, slope, and drift. (*Courtesy Bureau of Mines.*)

Fig. 17.7. Basic steps in conventional mining. (*Courtesy Bureau of Mines.*)

can be used where roof conditions are fair to poor. Strong roof rock, however, can be a problem. The seam should be over 42 inches thick to accommodate this type of operation and the large coal cutter or plow that is used. A large reserve is necessary.

A modification of this method, using a continuous mining machine on faces up to 150 feet long, is known as *sortwall mining*. It uses the room support system of self-advancing chocks developed for longwall operations.

Coal Preparation

"Coal preparation" is a term applied to upgrading of coal to make it suitable for a particular use. It includes blending and homogenization, size reduction, and beneficiation or cleaning. It is this last aspect, as well as the degree to which it is required, that most significantly governs the cost of coal preparation. Figure 17.9 shows levels of cleaning in terms of broad categories.

Only about one-third of the 3.3×10^9 metric tons of coal produced every year is at present cleaned by breaking, crushing, screening, and wet and dry concentrating processes. This section will briefly describe the processes utilized for coal cleaning.

Coal generally is cleaned by physical methods to remove mineral matter consisting of rock, slate, pyrite, and other impurities. In general, processes for the removal of mineral matter utilize either differences in density or surface chemical properties of coal and mineral matter. For convenience, coal cleaning can be divided into three parts: coarse-coal cleaning, medium-coal cleaning, and fine-coal cleaning. Figure 17.10 shows various coal cleaning methods and corresponding size ranges of coal treated.

The first step in most of the coal cleaning operations is size reduction, the main objective being to liberate mineral matter from coal. Size reduction equipment for coal application ranges from heavy-duty crushers and breakers, capable of crushing lumps of up to a meter cube in size, to coal pulverization equipment capable of milling coal to a fine powder. The hardness and the softness of the coal is defined in terms of the Hardgrove Grindability Index (HGI). A high (>80) HGI indicates soft coal,

Fig. 17.8. Continuous mining machine. (*Courtesy Bureau of Mines.*)

Fig. 17.9. Different levels of coal cleaning and their effect on coal recovery and economics.

easy to grind, and a low (~35) HGI indicates difficult-to-grind coal.

After the coal is crushed, it generally is screened to separate the raw coal into various sizes for cleaning operations. Screening of coal above 12 mm size usually is carried out dry. Double-deck vibrating screens commonly are used for this purpose. For sizing below 12 mm, wet screening is done, using either high frequency vibrating screens or the Sieve Bend. Details on size reduction can be found in various books on coal preparation.[7,8]

Fig. 17.10. Coal cleaning equipment in common use in the coal industry with respect to the coal size processed.

Coarse-Coal Cleaning

Jigging and dense-medium separation generally are used for coarse coal cleaning.

Jigging. Jigging is among the oldest of mineral processing methods and one of the first adopted for coal cleaning. The separation of coal from mineral matter is accomplished via a fluidized bed created by pulsing a column of water, which produces a stratifying effect on the raw coal. Water is pulsated through the perforated bottom of an open-top rectangular box. The lighter coal particles rise to the top, overflow at the end of the jig, and are removed as clean product. The denser mineral matter settles on a screen plate and is removed as refuse. Most jigs are of the Baum type (Fig. 17.11), in which the pulsations are generated by air pressure.

Fig. 17.11. Sections of a typical Baum jig.

Dense-Medium Washing. Dense-medium separations include processes that clean raw coal by immersing it in a fluid medium with a density intermediate between the density of clean coal and that of reject. Most dense-medium washers use a suspension of fine magnetite in water to achieve the desired density. The process is very effective in providing a sharp separation and is relatively low in capital and operating costs.

Medium Coal Cleaing

Medium-size coal includes coal ranging in size from 3/4 inch to 28 mesh (0.5 mm). The principal techniques utilized are the wet concentrating tables, the dense-medium cyclone, the hydrocyclone, and spirals. All of these equipment types are widely used in the coal industry.

Wet Concentrating Table. This is also known as a shaking table and works much like the classical miner's pan. The shape of the table is usually a parallelogram, and the deck of the table is sloped (or sloping) slope with rubber riffles on it. The upper side of the desk contains a feed box and a wash water box. The deck is vibrated longitudinally with a slow forward stroke and a rapid return. Normal decks are 5 meters long along the longest edge and 2.5 meters wide at right angles to the length. Such a machine can process up to 12 to 15 tons of coal/hour.

Dense-Medium Cyclone. The essential features of a cyclone are shown in Fig. 17.12. A mixture of raw coal and a dense medium (magnetite suspension) enters the cyclone tangentially near the top, producing free-vortex flow. The refuse moves downward along the wall and is discharged through the underflow orifice. The clean coal moves radially toward the cyclone axis, passes through the vortex finder to the overflow chamber, and is discharged from a tangential outlet.

Fig. 17.12. Heavy-medium cyclone separator.

Cyclones are available in various diameters; a 500-mm-diameter unit can process up to 50 tons of coal/hour, whereas one of 750 mm diameter can process up to 120 tons/hour.

Some other forms of the heavy-medium cyclone are the Vorsyl Separator, the Larcodems, the Dynawhirlpool, and the Tri-flo Separator. Each of these units utilizes cyclonging principles.

Hydrocyclone. When only water is used in a cyclone for cleaning coal, it is called a hydrocyclone. Hydrocycloning has been applied even to process coal finer than 0.5 mm. Its design differs from that of the dense-medium cyclone by having a greater cone angle and a longer vortex finder.

Spiral Separator. The spiral separator is usually 8 to 10 feet in height and consists of a trough going downward in a spiral. The coal slurry is fed in at the top, and as it follows the spiral down, centrifugal force separates the coal from the denser mineral matter. A splitter box at the bottom of the spiral can be adjusted to obtain the desired separation.

Fine Coal Cleaning

Froth Flotation. Coal below 0.5 mm in size, classified as fine coal, generally is processed by surface chemical methods such as froth flotation. In the froth flotation process, the fine coal slurry, to which a small amount of flotation agents—usually fuel oil and a short-chain alcohol (methyl isobutyl carbinol)—is added, is processed through a flotation cell. In the cell, fine bubbles are generated by using either forced air or suction. The coal, being hydrophobic, attaches to the air bubbles and rises to the top where it is removed as froth. The refuse, being hydrophilic, remains in the water and is removed from the bottom. The process is very effective in recovering high-grade coal at moderate cost.

The reagents used in froth flotation of coal are classified as collectors, frothers, and modifying agents. Collectors (kerosene or No. 2 fuel oil) provide a thin coating of oil on the coal and promote attachment of air bubbles to it. The frother (methyl isobutyl carbinol) promotes a stable froth. Modifying agents have several functions such as depressing unwanted material, altering the surface of the particles to aid attachment of air bubbles, and providing acidity ot alkalinity to the flotation pulp.

An advanced froth flotation technique called columm flotation has been shown to be superior in obtaining a low-ash clean coal with high recovery of combustible. It has been proved effective in recovering ultra-fine (minus 74 microns) coal.[9,10] A schematic diagram of the column flotation process is shown in Fig. 17.13.

The basic difference between the conventional and the column flotation system is in the design, which provides quiescent flotation conditions in the column, thus effectively recovering fine coal. Also, gentle washing of froth at the top of the column removes entrained impurities from the froth, providing a low-ash clean coal product. The colums also have been effective in removing high amounts of liberated pyritic sulfur. In 1989 the University of Kentucky Center for Applied Energy Research designed four 2.4-meter-diameter columns that were installed at the Powell Mountain Coal Company, St. Charles, Virginia, for recovery of fine coal from refuse streams.

Recently, an air sparged cyclone developed at the University of Utah was tested on a pilot scale. In this technique, flotation is performed inside a cyclone that has porous walls through which air is injected.

Fig. 17.13. Schematic diagram of column flotation.

Other Processes. An oil agglomeration process that utilizes oil or other hydrocarbon to agglomerate coal, leaving mineral matter in aqueous suspension, has been tested on pilot and commercial scales. The Otisca process utilizes pentane as an agglomerant, which is recovered from clean coal and recycled back in the process. The NRCC process from Canada utilizes fuel oil as an agglomerant. Another process, the high-gradient magnetic and electrostatic cleaning process, utilizes differences in magnetic and electrical charge properties of the mineral matter present on coal. Both of these process have achieved limited success.

Chemical cleaning processes utilize alkali or acid to leach the impurities present in the coal. TRW's Gravimelt process, which has been tested on a pilot scale, utilizes molten caustic to leach mineral matter. The major drawback of a chemical cleaning process is the economics.

The chemical methods are not effective in removing major amounts of organic sulfur. Thus, removal of total sulfur to produce a clean-burning fuel is very difficult. It is envisioned that the optimum desulfurization of coal will include physical, chemical, and microbial treatment. The raw coal will be ground fine to liberate most of the pyritic sulfur and then processed by using the column

flotation process to remove most of the liberated mineral matter. Then the clean coal will be treated chemically and finally by microbes to remove organic sulfur from the coal.

Fine Coal Drying

One of the biggest problems in processing fine-size coal is dewatering the drying. Conventional mechanical dewatering devices are effective in dewatering coarse and medium-size coal. For dewatering of fine-size coal, especially if a substantial amount is in the 74-micron range, the conventional techniques provide a dewatered product containing about 30 percent moisture.[11]

Filtration devices that utilize high pressure forces, such as Andaritz Hyperbolic and Ama's KDF filters, have the capability of reducing moisture in the filter cake to about 20 percent. However, the capital and operating costs are high for these devices. Other newly developed techniques that have been tested on a pilot scale include an ultrahigh-*g* centrifuge, which generates forces up to 4000 *g* in the centrifuge, and an electro-acoustic technique that utilizes the synergistic effect of electric, ultrasonic, mechanical, and surface chemical forces to remove moisture from the cake.

COAL UTILIZATION—PYROLYSIS

Technology and Relation to Other Coal Conversion Schemes

In a strict sense, coal pyrolysis means heating of coal in an inert atmosphere. When heated, coal undergoes thermal decomposition, which results in the evolution of gases and tars. In all coal conversion and utilization processes (including gasification, combustion, and liquefaction), some form of coal pyrolysis occurs. Because coal pyrolysis by itself is an important coal utilization process for making coke and oil, a vast amount of literature exists on this topic. By carefully varying the pyrolysis conditions and by adding reactive atmospheres (such as hydrogen), one can optimize production of a desired product (i.e., solid, liquid, or gas). Gasification is closely related to pyrolysis in that when coal is gasified by reacting it at a high temperature with steam and an oxidant (air or oxygen), pyrolysis is the first stage of the reaction as the coal is heated to the reaction temperature. Depending on the type of gasifier, the condensable hydrocarbon gases may be completely destroyed or may be collected as a by-product. Similarly, pyrolysis occurs at the first stage of combustion, but hydrocarbons are destroyed by oxidation. In direct liquefaction, some pyrolysis may occur because of the high temperature utilized when solvents are reacted with the coal, but the dominant process normally is a dissolution of the coal structure, due to solvent chemistry.

Mild gasification is a pyrolysis-based process designed to produce a product slate of alternative fuels by decomposing coal at relatively mild conditions of temperature and pressure. The idea is to produce the most profitable combination of products with a given feedstock and associated operating, distribution, and capital costs.[81,89]

History

Coal pyrolysis is a very old technique (dating back to the eighteenth century) that allows separation of the hydrocarbon fuels and uses relatively inelegant technology. Most pyrolysis systems in use in the late 1800s and early 1900s were in Europe and had as their objective the production of smokeless fuel (char) for domestic use. However, within a short time it was realized that the coal tar fraction contained valuable chemical products. For example, the olefins were described as having potential wealth similar to "Ali Baba's cave." Subsequently, interest faded because of the abundance of inexpensive petroleum. A detailed review has been published.[12]

Interest in pyrolysis was rekindled by the oil embargo of 1973, and emphasis was placed on developing processes that would achieve

a maximum yield of liquid products. Several processes were developed to the pilot-plant stage in the 1970s, only to be placed in mothballs when oil prices softened in the early 1980s.

Comparison of Pyrolysis Processes

Based on the reactor configuration, pyrolysis can be conveniently divided into two broad groups: advanced (contemporary) and early processes. A simplified classification of various pyrolysis procedures is provided in Table 17.4.

Coal pyrolysis processes generally are regarded as low temperature when the process temperature does not exceed 700°C, or high temperature if the temperature of the process is at or exceeds 900°C. Some processes operate at temperatures between these extremes. It needs to be emphasized that the pyrolysis conditions used determine the extent to which coal is carbonized and the nature of the liquid and gaseous products. A number of reviews on the advanced pyrolysis processes are available.[13,14] Only a brief comparison of various contemporary devolatilization processes is given in Table 17.5.

Factors Affecting Coal Pyrolysis

During pyrolysis, the yield of gases and liquid products may vary from 25 to 70 percent by weight. The yield depends on a number of variables, such as coal type, gas atmosphere, final pyrolysis temperature, time–temperature history, and pressure. A certain operating condition may lead to increased product yield, but achievement of this process condition often is obtained at added expense.

Depending on the desired products, one may choose to operate a given process in a particular time and temperature mode. Many excellent reviews (e.g., those of Howard,[15] Gavalas,[4] and Solomon and Serio[16]) have been published that discuss the factors affecting coal pyrolysis and product composition. In the following sections, major conclusions presented in the literature are discussed (as summarized in Table 17.6).

Temperature and Heating Rate. The effect of pyrolysis temperature on the yield of products is relatively straightforward. Coal undergoes a large variety of physical and chemical changes when heated gradually from ambient temperature to approximately 1000°C

TABLE 17.4 Simplified Classification of the Various Pyrolysis Procedures

Pyrolysis Process	*Final Temperature °C*	*Products*	*Processes*
		Early Schemes (Slow Heatup)	
Low temperature	500–700	Relatively more reactive coke.	Rexco (700°C; cylindrical units). Coalite (650°C) made in vertical tubes.
Medium temperature	700–900	Relatively more reactive coke with high gas yield, or domestic briquettes.	Town gas and gas coke (obsolete). Phurnacite, low-volatile coal, pitch-bound briquettes carbonized at 800°C.
High temperature	900–1050	Hard unreactive coke for metallurgical use.	Furnace coke (950–1050°C).
		Advanced Techniques (Rapid Heatup)	
Low temperature	500–700	Relatively higher tar yield (20–35 wt.) compared to gas (5–15 wt. %).	COED (fluid bed); Occidental flash (entrained flow) pyrolysis.
Medium temperature	700–900	Lower tar yield compared to that obtained in low-temperature operation.	Clean Coke Process; oxidized clean coal (in fluid bed).

TABLE 17.5 Representative Operating Conditions and Yields for Contemporary Coal Pyrolysis, Hydropyrolysis, and Extraction Processes

Process	*Status*	*Reactor Type*	*Reactor Temperature °(C)*	*Reactor Pressure (MPa)*	*Residence Time (s)*	*Coal*	*Yield, Mass Percent Dry Coal*				*Size of Facility (in 1000 kg/day Coal)*
							Char	*Tar/Oil*	*Gas*	*Water*	
COED, FMC Corporation	Developed	Multiple fluidized-bed	290–565	0.12–0.19	1200[1]	Bituminous	62	21	14	3	32
TOSCOAL	Developed	Rotating horizontal retort	520	0.1	300–600	Subbituminous	69	13	9	9	23
Lurghi-Ruhrgas	Commercial	Mechanical (twin screw) mixer	595	0.1	~3	Subbituminous	50	~32[2]	11	~7[2]	1600
Occidental Flash Pyrolysis	Developing	Entrained flow	610	0.3	1.5	Bituminous	56	35	7	2	3
Clean Coke, U.S. Steel	Developing	Fluidized-bed	~800	~0.8	~3000	—	66	14	15	—	—
Rockwell/Cities Service Flash Hydropyrolysis	Developing	Entrained flow	845	3.5	~0.1	Bituminous	46	38	16	—	22
Supercritical Gas Extraction, NCB[3]	Developing		400	10	1800	Bituminous	63	33[4]	2	2	—

[1]In pyrolysis stages.
[2]Water plus tar/oil sum 39%, individual values estimated.
[3]NCB: National Coal Board, United Kingdom.
[4]Tar extract, softening point 70°C.
Source: Adapted from Probstein and Hicks.[13]

TABLE 17.6 Summary of Advantages and Limitations of Pyrolysis Process Conditions and Feedstock Characteristics

Heating Rate

- High heating rate increases liquid/gas yield and reduces char yield.
- The tar obtained at a high heatup rate is of poorer quality (i.e., lower H/C ratio) than that obtained at a slower heating rate.[13,15]
- High heating rate increases open char structure and char reactivity (in reactive gases).
- Sophisticated (often exensive) systems needed to achieve high heating rate.
- High heating rate increases the thermoplastic (softening and swelling) behavior of coal.

Temperature

- Low-temperature operation (500–700°C) improves liquid yield.
- Affects heteroatom distribution among char, liquid, and gas.
- At elevated temperatures (> 1,300°C), inorganics are removed as slag.
- Longer residence time needed for reaction to be completed at lower temperatures.

Pressure

Inert Gas Atmosphere:

- Higher pressure operation reduces reactor size needed (i.e., increases throughput).
- Higher pressure reduces tar yield.
- Coal feeding, product separation are more difficult at high pressure.
- Better gas-solid heat transfer is achievable at higher pressure.

H_2 Atmosphere:

- Cost of H_2 must be considered versus the quality of the product generated.
- Improves yields of liquid and light products.
- Requires sophisticated pressure systems.
- May increase the undesirable agglomerating properties of coal.[17]

Other Atmospheres (H_2O, CO_2, CO, CH_4, CS_2):

- Probably improve liquid/gas yield.
- Not much information is available.

Vacuum:

- Plastic behavior of coal is reduced.[17]
- Increases liquid/gas yield.[17]
- It is difficult to achieve gas–solid heat transfer (solid–solid heat transfer feasible).
- Not much information is available.

Particle Size

- Smaller particle size improves product (gas/liquid) yield.
- Smaller particle size reduces secondary reactions.
- Grinding cost increases with the reduction in size.

Coal Rank

- High Volatile A (HVA) bituminous coals produce largest quantities of tar.
- Lignites are rich in oxygen functional groups which lead to overall reduction in the calorific value of the product.
- Types of sulfur (pyritic versus organic) present in various coals can influence its distribution in products.

Some changes are observed before the inset of thermal decomposition (i.e., carbon–carbon bond scission and the like), which occurs above 350°C. When heated to approximately 100°C, the physically sorbed moisture is liberated. For low-rank coals such as lignites that contain appreciable carboxylic functions as part of the coal structure, heating will liberate carbon dioxide by thermal decarboxylation. Over 50 percent of the carboxylic acid

functions are lost as carbon oxides in the temperature range of 100 to 250°C. As the temperature of thermal treatment increases to the range of 200 to 400°C, coals may lose a number of lower molecular weight organic species (especially aliphatic compounds), which are believed to arise from moieties (components) that are loosely bound to the more thermally stable part of the coal structure. At a higher temperature (375–700°C), depending on the heating rate, thermal destruction of the coal structure occurs, which is reflected by the formation of a variety of hydrocarbons including methane (and other alkanes), polycyclic aromatics, phenols, and nitrogen-containing compounds. In this temperature range, bituminous coals soften and become plastic (thermoplastic) to varying degrees. At a still higher temperature (600–800°C), depending on the heating rate, the plastic mass undergoes repolymerization reactions, forming semicoke (solid coke still containing significant volatile matter). At temperatures exceeding 600°C, semicoke hardens to form coke with the evolution of methane, hydrogen, and traces of carbon oxides. Pyrolysis of coal is essentially complete at approximately 1000°C. The temperature at which the maximum devolatilization rate occurs depends on the heating rate. For a slow heating rate (about 5°C/sec), the maximum occurs around 400°C, whereas for a rapid heating rate (about 10°C/sec), it might occur at 900°C.[18] Soaking coal at the final temperature normally would increase the yield because decomposition of char is a relatively slow process.

Since the changes and reactions described above do not occur at sharply defined temperatures, it is to be expected that heating rate will have a marked effect on the nature and distribution of pyrolysis products resulting from a given process. The general situation is as shown below:

Temperature	*Changes/Reactions*
100°C	Physically sorbed water is liberated.
100–250°C	Thermal decarboxylation of low-rank coals liberates CO_2.
200–400°C	Lower molecular weight organic species, especially aliphatic compounds, are lost.
375–700°C	Thermal destruction of coal structure occurs; there is formation of methane and other alkanes, polycyclic aromatics, phenols, and nitrogen-containing compounds.
1000°C	Pyrolysis is essentially complete.

Solomon and coworkers conducted extensive work on the kinetics of coal devolatilization, and many reviews are available.[19]

Coal Type. The type of coal utilized strongly influences its pyrolysis behavior. Low-rank coals (e.g., lignites) contain oxygen functional groups that evolve upon pyrolysis, primarily as water and oxides of carbon. The high-rank bituminous coals contain lesser amounts of oxygen; consequently, these coals produce significantly lower volumes of water and carbon oxides than hydrocarbon gases (primarily tars) when pyrolized. The nature of the tar produced also is dependent on coal rank. Bituminous coal tars tend to be more aromatic in nature (and relatively more thermally stable) than the tars generated from lignites. The type of coal pyrolyzed also determines its physical behavior when pyrolyzed. When heated, bituminous coals soften, become plastic, and swell to varying degrees. Lower-rank coals generally do not become plastic. However, at rapid heating rates (about 10°C/min) or at elevated gas pressure, certain lower-rank coals may melt and demonstrate some plastic and swelling characteristics.

Other Factors. Several other factors influence the nature of the pyrolysis process and the direction it takes, some to a lesser extent and others to a greater extent. These include the size of the coal particles, size and configuration of the coal bed, pressure (or vacuum) under which the process is operated, presence of various inorganic materials in the coal, secondary reactions, and others.[86] Although it is beyond the scope of this chapter to consider these items, interested readers can find additional information in the literature. These sources include reports on pressure effects,[15,20] effects of vacuum,[17] effects of

inorganics,[21,86] and a potentially reactive gas atmosphere.[17,22]

Pyrolysis Products: Utilization and Characterization

Efficient utilization of all generated solid, liquid, and gaseous products is vital for the overall process economics of a pyrolysis system. A number of viable alternatives can be considered for utilization of these products, some of which are discussed below. The generated products may require varying degrees of treatment before they become usable. For example, the stream exiting the pyrolyzer unit will require separation of gas, liquid, and particulates. The unit operations (e.g., hot dust removal cyclones, quench/particulate scrubber towers, and venturi scrubbers to remove tar mist) required for separation may be similar to those used for a fixed-bed gasifier (e.g., cold gas cleanup). The gas stream may require, depending on its end usage, a sulfur removal process. In addition, a pyrolysis plant also will require processes for wastewater treatment. Finally, environmental factors such as toxicology of the coal pyrolysis liquids (CPL) need to be considered. Mutagenicities of CPL are strongly dependent on the conditions of pyrolysis (e.g., temperature, coal type, and gas atmosphere).

Tar Utilization. Markets tend to have product specifications of various liquid fuels that do not vary widely. However, the characteristics of CPL can vary greatly, depending on the coal type and the operational conditions chosen. In general, the H/C (atomic) ratio for CPL is lower than that for various petroleum feedstocks. This ratio varies from 1.8 to 1.9 for No. 2 fuel oil, and from 1.7 to 1.8 for No. 6 fuel oil. In comparison, the H/C (atomic) ratio for CPL ranges from 0.9 to 1.5. In addition, CPL appear to be significantly more aromatic (e.g., 50–70%) than petroleum-based fuels.

The uses of CPL can be crudely divided, for convenience, into two broad categories: (a) direct combustion applications (uses that do not require extensive upgrading), and (b) transportation fuels and chemicals (uses requiring upgrading, extensive processing, or both). Because of the high aromaticity of CPL, one may consider converting them into gasoline (high proportion of branched-chain alkanes and lighter aromatics) rather than diesel (straight-chain alkanes with heavier aromatics). A lot of attention in low-temperature tar processing was devoted to hydrogenative techniques (hydrotreating, hydrocracking, or both). The primary objectives have been to: (a) reduce viscosity, (b) reduce the polynuclear aromatics content, and (c) remove heteroatoms (sulfur, nitrogen, and oxygen) in converting these tars into usable fuels and chemicals.

The cost of H_2 still is the primary impediment in tar upgrading. Low-temperature tars can be used as chemicals: phenols and phenolic chemicals, road tars, preservatives, and carbon binders, to name a few. These usages, however, do not constitute a large-enough market in themselves to support a major industry.

Char Utilization. Char produced in the pyrolysis process can have a range of applications to offset the cost of producing liquid fuels. One obvious application is on-site combustion to generate heat for the pyrolysis process itself. Alternatively, chars can be used for combustion in electric power plants. Char can also be gasified to produce hydrogen for use in the pyrolysis plant if the process requires hydrotreatment (e.g., tar upgrading or hydropyrolysis). Other applications include production of activated carbon and molecular sieves.

In general, the volatile matter content of char is significantly lower than that of the parent coal. It was demonstrated, however, that the reactivity of several low-temperature chars (prepared at 500°C) was greater than that of the parent coal.[23]

Economic and Process Considerations

Using current coal technology, coal pyrolysis can produce liquid fuels that are potentially directly usable as fuels, or fuels that can be upgraded for such uses. However, it does not

appear to be economical at present to produce solely liquid fuels as a substitute for their petroleum counterparts. Plant designs for coal pyrolysis projects aimed solely at this purpose generally require more sophisticated processing steps than their petrochemical counterparts to produce liquid hydrocarbons that can meet required specifications, while at the same time minimizing waste production. Steps involving upgrading and refining of tars by hydrogenation and hydrocracking to produce premium fuels require massive amounts of expensive hydrogen. These costly endeavors can be justified only in situations where assurances of ample supplies will take precedence over any economic considerations, as was the case in World War II Germany.

The economics of a coal pyrolysis process to cream-off the liquids, utilizing mild conditions (less sophisticated and thus less expensive systems than direct or indirect liquefaction processes) before char is combusted or gasified, can be quite attractive. A number of process streams can be considered for efficient use of all generated solid, liquid, and gaseous products. For example, chars can be sold to offset the production costs of pyrolysis liquids. The gases produced can be combusted to supply the heat generated for the pyrolysis process.

COAL UTILIZATION-GASIFICATION

History

Coal gasification is an old, established technology that has been rediscovered and is gaining favorable interest. A large number of coal gasifiers operated commercially in this country and in Europe to produce industrial and residential fuel gas during the last and the early part of this century. By the mid-1950s, however, the availability of inexpensive natural gas and petroleum had led to the abandonment of most coal gasification processes.[93]

In the years since 1973, the price of petroleum has increased. In addition, the capability of integrated gasification combined cycle power plants to effectively compete with direct coal combustion for electric power generation has significantly increased the attractiveness of coal gasification.

Coal was first gasified in England by William Murdock in 1792, and the world's first coal gas company was chartered in England in 1812. Coal gas was first produced in America in 1816 in the city of Baltimore. However, by 1850, over 55 commercial coal gasification plants were generating gas for lighting and heating in the United States. In the early 1930s there were over 11,000 coal gasifiers operating in the United States, consuming about 12 million tons of coal and producing 500 million Btu/day of coal gas. Most of the early gasifiers were moving-bed units operating with sized coal charged with steam and air to produce a "product gas" (150 Btu/scf). Operation of the moving-bed gasifier in a cyclic mode (blowing the gasifier first with air, followed by steam to gasify) increased the heating value of the product "water gas" to 300 Btu/scf. Coke oven gas was used as "town gas," while co-firing of oil with water in the moving bed process produced "carburetted water gas."

Coal gasification plants continued to be built on a limited scale in certain areas of the world including the United States, Europe, and countries with a limited supply of petroleum. Extensive large-scale development work was done in the United States in the late 1940s to mid-1950s, including: (a) development of a 100 ton/day Texaco entrained-flow gasifier at Morgantown, West Virginia, by Texaco and Olin Matheson; (b) demonstration of a 360 ton/day Babcock and Wilcox entrained gasifier at Belle, West Virginia; (c) development of a 24 ton/day Koppers-Totzek entrained gasifier at Louisiana, Missouri.

During the 1950s and 1960s a significant contribution to domestic energy use was made by coke oven gas, which is roughly similar to the gas generated in moving-bed gasifiers.

New interest in coal gasification since 1973 has led to a number of large-scale demonstration projects for the following

processes:

- Texaco
- British Gas Lurgi/Slagger
- Kilngas
- Shell and Dow gasifiers

Development of these coal gasification processes assures the potential for coal gasification technology.[94]

Reaction Fundamentals

The initial steps in a coal gasification process involve grinding and/or pretreatment of the coal. In a gasifier, the coal is heated in the presence of reactive gases whose composition depends on the product desired. The first event that occurs in the gasifier is pyrolysis or devolatilzation. Walker[24] has presented extensive reviews of gasification reaction fundamentals.

During gasification, coal undergoes some or all of the following reactions/changes:

- Pyrolysis (devolatilization)
- Char–oxygen reaction
- Char–CO_2 reaction
- Char–H_2O reaction
- Char–H_2 reaction

In the manufacture of synthetic natural gas, all of the above reactions occur to varying degrees; however, in synthesis gas (CO/H_2 mixtures) production, the char–H_2 reaction is suppressed.

Coal char can resemble graphite, depending on the conditions under which it is formed. Coal-char gasification systems generally are bounded by equilibria obtained by considering the carbonaceous solids to be equivalent to graphite. Therefore, it is common to consider char as similar to graphite in a thermodynamic sense.

Char formed in pyrolysis has a higher degree of crystallinity than the starting coal but is significantly less ordered, in terms of crystallinity, than graphite. In the case of the latter, the carbon atoms located at the edges of crystallites have unpaired sigma electrons that are most susceptible to reaction. The concentration of active edge sites depends on crystallite size and concentration of defects such as vacancies and dislocations. In the case of coal chars, the concentration of defects depends on the coal type and pyrolysis conditions. In general, bituminous coal chars have lower active site concentration and thus lower reactivity.

Heteroactoms such as O, N, and S can play a role during gasification. Hydrogen sites are presumed to increase char reactivity by preferential oxidation with subsequent production of nascent carbon active sites of high activity. Nitrogen and sulfur also could encourage ring attack owing to concentration of pi electrons.

The pore structure can markedly affect char reactivity. Most coals in general, and coal chars in particular, are highly porous and contain a polymodal pore size distribution. Pores normally are classified into macropores (>500 Å in diameter), transitional pores (20–500 Å in diameter), and micropores (<20 Å in diameter). Upon pyrolysis, pores in coals open up but still contain microporosity. Coal chars, in general, and lignitic chars, in particular, retain the polymodal pore distribution. The surface areas of coal chars can range between 100 and 800 m^2/g. Most of this surface area and therefore the active surface area resides inside the char particles; so the accessibility of the reactive gases to the active sites is very important.

The fundamental reactions involved during gasification can be described by the Langmuir-Hinchelwood mechanism and the Rideal-Eley mechanism. The Langmuir-Henshelwood mechanism involves three different steps: (a) adsorption of the gas onto the solid surface; (b) surface reaction/migration; (c) desorption of the product species from the solid surface. In the Rideal-Eley mechanism, the basic steps involved are: (a) reaction between the gas molecular and surface atoms by direct collision; (b) desorption of formed product(s).

These mechanisms describe the intrinsic reaction rate, which would be the upper limit of the achievable gasification rate. The actual

gasification rate in a gasifier would be equal to or lower than the intrinsic rate. Determination of the intrinsic reaction rate for pure carbon, in general, and coal chars, in particular, currently is not possible because nearly all coal chars and most carbons are ill-defined and/or impure.

Because coal chars are highly microporous, most of the gasification reaction take place inside the char particles. Therefore, they require diffusion of gas into and diffusion of products out of the porous particles. The overall diffusion process can be described by the following steps: (a) diffusion of reactant from the bulk gas to the solid surface (film diffusion); (b) diffusion of reactant from the surface to the interior (internal diffusion); (c) diffusion of product from the interior to the surface (internal diffusion); (d) diffusion of product from the surface to the bulk gas (film diffusion).

Film diffusion coefficients normally are large and have very low temperature dependency ($T^{0.5-1.5}$). Internal diffusion coefficients can be activated (for pores 5–7 Å in diameter with E_{act} of about 5 to 10 kcal) or nonactivated. For pores in the transitional size range, internal diffusion is of the Knudsen type with a low temperature dependency ($T^{0.5}$).

One can calculate the rates of internal and film diffusion if the pore diameter and process conditions are well defined. The temperature dependency of the rate can be presented in the form of an Arrhenius plot (i.e., log rate versus reciprocal temperature). Gasification rates can be divided into three zones, I, II, III, depending on various flow and diffusion condition. To determine the overall gasification rate in a gasifier when gas diffusion controls the overall rate, it is necessary to model the actual gasifier.

C–O_2 Reaction. This is the most common and the fastest gasification reaction and is highly exothermic. The primary reaction products, as shown below, are CO and CO_2:

$$C_s + O_2 = CO_2 \qquad -94.1 \text{ kcal/gmole}$$

$$C_s + 0.5O_2 = CO \qquad -26.45 \text{ kcal/gmole}$$

The equilibrium constants, K_p, for both reactions are >1 up to 3500°C, and, as a result, both reactions are essentially irreversible up to 3500°C. The CO/CO_2 should increase with lower pressure and high temperature, as pressure does not influence the equilibrium of the first reaction, and the forward reaction (second reaction) is not favored under pressure.

The postulated elementary reactions for the carbon–oxygen reaction are:

$$2C_f + O_2 \rightarrow 2C_f(O)$$

$$C_f(O) \leftrightarrow C_{fs}(O)$$

$$C_f(O) \rightarrow CO$$

$$C_f(O) + C_f(O) \rightarrow CO_2 + C_f$$

In the above, C_f stands for a free active site, $C_{fs}(O)$ represents a stable oxygen surface complex, and $C_f(O)$ is a transient oxygen surface complex.

Above 2300°K the reaction gas dissociates very quickly, but above 1700°K the active sites are quickly annealed. Thus the reaction rate increases with temperature (the exception being between 1700 and 2300°K).

C–H_2O Reactions. The carbon-steam reaction is an endothermic reversible reaction:

$$C_{(s)} + H_2O \leftrightarrow CO + H_2 \qquad 31.4 \text{ kcal/gmole}$$

Mechanism I:

$$H_2O + C_f \leftrightarrow H_2 + C(O)$$

$$C(O) \rightarrow CO + C_f$$

Mechanism II:

$$H_2O + C_f \rightarrow H_2 + C(O)$$

$$C(O) \rightarrow CO + C_f$$

$$C_f + H_2 \rightarrow C_f(H_2)$$

The following presents the rate expression based on either mechanism:

$$\text{Rate} = \frac{A \cdot P_{H_2O}}{1 + B \cdot P_{H_2} + C \cdot P_{H_2O}}$$

In the above expression, A, B, and C are constants, and P_{H_2O} and P_{H_2} are partial pressures of H_2O and H_2. The reaction rate is not favored by high pressure. Both H_2 and H_2O inhibit the reaction.

The carbon–steam reaction has a side reaction, called the water–gas shift reaction,

shown below:

$$CO + H_2O = CO_2 + H_2$$

This reaction, which is very rapid, is catalyzed by various impurities and surfaces.

CO_2–Solid Reactions. The equilibrium for the carbon-CO_2 reaction is favored at high temperatures and at lower pressures:

$$C_2 + CO_2 = 2CO(g) \qquad 41.2\,\text{kcal/mol}$$

Two mechanisms have been proposed for this reaction:

Mechanism I:

$$CO_2 + C_f \leftrightarrow CO + C(O)$$
$$C(O) \rightarrow CO + C_f$$

Mechanism II:

$$CO_2 + 2C_f \rightarrow C(CO) + C(O)$$
$$C(O) \rightarrow CO + C_f; \quad C(O) \leftrightarrow C_f + CO$$

In mechanism II, CO inhibits by adsorbing onto the active sites. In mechanism I, CO does not adsorb on the active sites, but the reaction is inhibited by CO because of its reaction with the oxygen complex.

Both reaction mechanisms yield a similar kinetic expression for the reaction:

$$\text{Rate} = \frac{A \cdot P_{CO_2}}{1 + B \cdot P_{CO} + C \cdot P_{CO_2}}$$

In the above expression, A, B, and C are constants, and P_{CO} and P_{CO_2} are partial pressures of CO and CO_2. The reaction rate becomes significant above 900°K. Chemisorption of CO_2 on carbon occurs even at 625°K. H_2 strongly inhibits the reaction, presumably by adsorbing on the active sites.

Types of Coal Gasifiers

All gasification processes can be separated into three generic types of reactors: (a) moving-bed or countercurrent reactors; (b) fluidized-bed or back-mixed reactors; (c) entrained-flow reactors (not back-mixed).

Figure 17.14 shows the generic types of gasification reactors together with temperature profiles and locations of reactants and products. Table 17.7 summarizes the important characteristics of each type of gasification process, Table 17.8 presents the overall status of various coal gasification technologies, and Table 17.9 presents the performance characteristics of selected gasifiers on Illinois No. 6, Pittsburgh No. 8, and West Kent coals.

Moving-Bed Gasifiers. Moving-bed gasifiers consist of a downward-moving bed of coal being contacted by upward-flowing gases that move through the bed. The coal sequentially undergoes drying, devolatilization, gasification, combustion, and ash cooling.

In the drying zone at the top of the bed, coal is dried by the upward-flowing hot gases. As the coal moves farther down the bed, its temperature increases, and at about 600 to 900°F pyrolysis occurs, liberating tars/oils and gases. The devolatilized coal, or char, is gasified by reaction with steam, carbon dioxide, and hydrogen to produce a mixture of carbon monoxide, hydrogen, and methane, along with unreacted steam and carbon dioxide, as char descends farther through the gasifier. The final composition of the gases produced and the rates of the reactions are strongly influenced by the gasifier temperatures and pressure of operation, as discussed above.

The peak temperature within the gasifier (occurs in the combustion zone) is controlled by the relative proportions of steam and oxygen (air) added to the gasifier. If dry ash is desired, sufficient steam is added to offset the exothermic oxidation reactions with endothermic steam–carbon reactions and to maintain temperatures in the combustion zone safely below the ash-fusion temperature. Slagging gasifiers operate at higher temperatures and remove ash in a molten state.

Moving-bed gasifiers require sized coal for proper operation; typically coal between $\frac{1}{4}$ inch and 2 inches in diameter is used. The countercurrent flow of gases and solids leads to high thermal efficiencies, with exit gas temperatures usually in the range of 500 to 1000°F. The raw product gas contains fines

Fig. 17.14. Generic coal gasification reactors. (*Source: Electric Power Research Institute and Synthetic Fuels Associates, Inc.*)

TABLE 17.7 Features of Fixed-Bed, Fluidized-Bed, and Entrained-Bed Gasifiers

PARAMETERS		
Moving Bed	*Fluidized Bed*	*Entrained Bed*
Gravitating moving bed of coal.	Fluidized beds arranged in one or more zones.	Up flow or down flow suspension gasifications.
Mechanical grates/distributors.	Uniform temperature and compositions throughout each fluidized zone.	High temperature–high rate process.
Discrete zones: • Preheating-drying devolatilization • Gasification • Combustion	Moderate process response.	Fast process response.
ADVANTAGES		
Moving Bed	*Fluidized Bed*	*Entrained Bed*
High carbon conversion efficiency.	High degree of process uniformity.	Handles all types of coal—no pretreatment.
Low ash carryover.	Excellent solids/gas contact.	Low steam consumption.
Low-temperature operation.	Lower residence time than fixed bed gasifier.	Excellent solids/gas contact.
Lowest air/oxygen requirement.	Higher coal throughput per unit volume of reactor.	No tar formation.
		No phenol formation.
		Ability to slag ash.
		High capacity per unit volume of reactor.
		Produces inert slagged ash.
DISADVANTAGES		
Moving Bed	*Fluidized Bed*	*Entrained Bed*
Sized coal required.	Sized coal required.	Requires finely crushed coal.
Coal fines must be briquetted.	Dry coal required for feeding.	Small surge capacity requiring close control.
Low offgas temperature.	Requires complicated gas distributor.	

and tar/oil produced during gasification. These materials are removed prior to downstream processing.

Moving-bed gasifiers can be operated at atmospheric pressure or can be pressurized: with either air or oxygen as the oxidant; with either dry ash or molten ash removal; with or without stirrers to prevent agglomeration.

Figure 17.15 illustrates the mechanical configuration for a Lurgi gasifier, and Table 17.10 shows the characteristics of several moving-bed gasifiers.

Fluidized-Bed Gasifiers. In a fluidized bed gasifier, reactant gases are introduced through a distributor at the bottom of the bed at velocities sufficient to suspend the incoming particles. The result is a bed of highly mixed solids in intimate contact with the gas phase. The agitation leads to uniform temperatures throughout the bed, and the high degree of mixing results in reaction rates that are generally higher than those experienced in moving-bed gasifiers. Fluidized-bed gasifiers generally do not achieve the high thermal efficiencies associated with moving-bed gasifiers.

Exit gas temperatures for fluidized-bed gasifiers usually are higher than those for moving-bed gasifiers. Therefore, the product

TABLE 17.8 Status of Various Gasification Technologies

Gasifier	*Project or Company*	*Location*	*Coal Throughput, tonnes/day*	*Product (final)*	*Start-up*
2nd generation gasification plants in commercial operation					
Texaco	Tennessee Eastman	Tennessee, USA	820	methanol	1983
Texaco	Cool Water	California, USA	910	IGCC	1984
Texaco	Ube Industry	Ube, Japan	1500	ammonia	1984
Texaco	Ruhrkohle/Ruhrchemi	Oberhausen, FRG	1720	oxo chemicals	1986
HTW[1]	Kemira Oy	Oulu, Finland	600	ammonia	1988
2nd generation gasification plants in demonstration operation					
Ruhr 100	Ruhrgas	FRG	100–200	fuel & synthesis	1979
BG/Lurgi	British Gas	Westfield, UK	500	fuel & synthesis	1984
Kiln Gas	Allis-Chalmers	Illinois, USA	540	fuel gas	1983
U-Gas	IGT	Chicago, USA	27	fuel & synthesis	1974
KRW[2]	Kellog-Rust-Westinghouse	Pennsylvania, USA	14–27	fuel gas	1975
HTW	Rheinbraun	Berrenrath, FRG	720	methanol	1985
VEW[3]	VEW	Dortmund, FRG	240	fuel & synthesis	1985
PRENFLOW[4]	Koppers-Totzek	Furstenhausen, FRG	48	fuel & synthesis	1986
Shell	Shell	Texas, USA	230–260	fuel & synthesis	1987
Dow	Dow	Louisiana, USA	2200	IGCC	1987
Saarberg-Otto	Saarberg-Otto	Furstenhausen, FRG	260	fuel & synthesis	1979
MIP[5]	Humboldt/Sumitomo (MEFOS)	Lulea, Sweden	240	fuel & synthesis	1986
2nd generation gasification plants under construction or planned					
NEDO[6]	NEDO	Chiba, Japan	20–50	hydrogen	1991
NEDO	NEDO	Iwaki, Japan	200	IGCC	1991
KRW	Appalachian project	Pennsylvania, USA	440	IGCC	1990
U-Gas	CCC-FWPSC	Virginia, USA	450	IGCC	1990

Acronyms:
1. HTW: High Temperature Winkler.
2. KRW: Kellogg-Rust-Westinghouse.
3. VEW: Vereinigte Elektrizitatswerke.
4. PRENFLO: Pressurized Entrained Flow Gasification.
5. MIP: Molten Iron Pure Gas.
6. NEDO:New Energy and Industrial Development Organization Japan.

Source: Takematsu, T., and Maude, C., "Coal Gasification for IGCC Power Generation," International Energy Agency, IEA CR/37, Mar. 1991.

gas contains much lower levels of tar/oil produced during pyrolysis. Particle attrition in the gasifier, however, generally leads to higher levels of fine/dust carryover. Particles and ash are removed in fluidized-bed systems as dry solids. Pretreatment of coal is required to handle strongly caking coals.

Fluidized-bed gasifiers can be operated at atmospheric pressure or pressurized. Fluidizing gases can be mixtures of steam with either air or oxygen and, in some cases, mixtures of hydrogen with other gases.

Table 17.11 summarizes the characteristics of several fluidized-bed gasifiers presently under development, as well as the Winkler gasifier, which is commercially available. There are several Winkler gasifiers at various installations worldwide, and the largest unit processes 650 tons of coal/day. The Winkler gasifier is illustrated in Fig. 17.16.

Entrained-Flow Gasifiers. Entrained-flow gasifiers consist of finely ground coal entrained in a flow of reactant gases. Contact between

TABLE 17.9 Performance Characteristics of Selected Gasifiers on Illinois No. 6, Pittsburgh No. 8, and West Kent Coals*

Parameters	*Lurgi dry ash*	*Lurgi slagger*	*Texaco*	*Shell*	*KRW (PDU, TP-034-2)*	*KRW design estimate*	*U-Gas West Kent coal*
	Illinois No. 6 coal				*Pittsburgh No. 8 coal*		
T, °F	1600	2700	2700	3000	1793	1850	1831
p, psia	315	300	600	365	230	600	15
H_2O/coal, mole/mole	2.42	0.298	—	0.027	0.43	0.29	1.34
O_2/coal, mole/mole	0.286	0.259	0.463	0.435	0.44	0.30	0.45
C conversion, %	99.3	99.5	99	99.3	76.0	89.6	93.8
Gas composition (dry)							
CO	15.36	58.05	51.69	61.46	43.38	51.5	26.1
CO_2	31.4	1.94	10.6	1.65	35.30	9.3	23.4
H_2	42.9	30.41	35.1	30.6	18.04	25.9	37.4
CH_4	8.78	7.76	0.09	0.04	1.83	10.4	2.8

*Shinner, R., Fourth Technical Meeting of COGARN, presented in "Coal Gasification: Direct Applications and Synthesis of Chemicals and Fuels: A Research Needs Assessment," by Penner, S. S., and others for US DOE.[25]

Fig. 17.15. Typical Lurgi gasifier. (*"Comparative Study of Coal Gasification Processes—Koppers-Totzek, Lurgi and Winkler," Koppers Co., Pittsburgh, PA.*)

TABLE 17.10 Characteristics of Moving-Bed Gasifiers

	Operating Pressure	*Oxygen or Air*	*Agglomeration Prevention*	*Status*
		Dry Ash, Single Stage		
GEGAS	To 500 psig	Air	Stirrer paddles	5-ft-diameter test unit
Lurgi	To 450 psi	Oxygen or air	Rotating blades	Commercial
MERC	to 105 psi	Air	Spiraling stirrer	42-in.-diameter test unit
Riley-Morgan	40 in. H_2O	Air	Agitator in rotating bed	10.5-ft test unit
Wellman-Galusha	10 in. H_2O	Air	Spiraling arms	Commercial
Wilputte	Atm	Air	Rotating arm	Commercial
		Dry Ash, 2-Stage		
ATC/Wellman (Incandescent)	Atm	Air	None	Commercial
FW/Stoic	Atm		None	Commercial
Ruhr-100	1500 psi	Oxygen	Stirrer blades	3–7 ton/hr test unit
Woodall-Duckham	40 in. H_2O	Air or oxygen	None	Commercial
		Slagging		
BGC/Lurgi	To 400 psi	Oxygen	Stirrer	6-ft-diameter test unit
GFERC	To 400 psi	Oxygen	Stirrer	16-in.-diameter test unit

Source: Bodle, W. W., and Schora, F. C., "Coal Gasification Technology Overview," *Symposium Papers: Advances in Coal Utilization Technology*, pp. 11–34, Institute of Gas Technology, Louisville, KY, May 14–18, 1979.

TABLE 17.11 Characteristics of Fluidized-Bed Gasifiers

	Developer	*Oxygen or Air*	*Pressure*
Winkler	Davy Powergas	Oxygen or Air	ATM
Rheinbraun	Rheinische Braunkohlenwerke	Oxygen	150 PSI
CO_2 Acceptor	Conoco Development	Air	150 PSI
Hygas	IGT/GRI/DOE	Oxygen or Air	1200 PSI
Synthane	DOE/PETC	Oxygen	1000 PSI
Westinghouse	Westinghouse Electric/DOE	Air	225 PSI
U-Gas	IGT/DOE	Oxygen or Air	350 PSI
Cogas	Cogas Development Co.	Air	10 PSIG
Catalytic	Exxon	None	500 PSI

Source: Bodle, W. W., and Schora, F. C., "Coal Gasification Technology Overview," *Symposium Papers: Advances in Coal Utilization Technology*, pp. 11–34, Institute of Gas Technology, Louisville, KY, May 14–18, 1979.

the solids and gases is cocurrent, with little or no batch-mixing occurring. This type of gasifier may be either single-stage or two-stage.

In general, high temperatures, 2200 to 3200°F, are used to completely gasify the coal in mixtures of steam and oxygen (air). Because of the high temperatures employed, gasification rates are considerably higher than those achieved with either fixed-bed or fluidized-bed gasifiers. These gasifiers can handle all coals, including those that are strongly caking, without pretreatment. The high temperatures of operation produce a gas devoid of both methane and tars/oils. Two-stage gasifiers are those in which the incoming coal first is entrained with reactant gases to produce gas; the resultant char is gasified further in a second stage, which may or may not be entrained. As is the case with fluidized-bed

Fig. 17.16. A Winkler gasifier. (*"Comparative Study of Coal Gasification Processes—Koppers-Totzek, Lurgi and Winkler," Koppers Co., Pittsburgh, PA.*)

gasifiers, stage operation enables better overall thermal efficiencies to be achieved without sacrificing higher gasifier throughputs, as the more reactive incoming coal can be gasified at lower temperatures than the less reactive chars. Entrained-flow gasifiers can be operated at atmospheric pressure or pressurized, and ash may be removed either dry or molten.

GASIFICATION FOR POWER GENERATION

Numerous studies confirm that gasifiers coupled with gas turbine/steam turbine combined cycles represent one of the most promising technologies for future coal-based power generation systems. Integrated Gasification Combined-Cycle (IGCC) technologies offer the potential for high electrical conversion efficiencies and economies of scale, which can be achieved in sizes below 1000 MW.

Compared to coal-combustion plants, IGCCs have lower capital and operating costs, resulting in lower net electricity costs. High reliability and low capital/operating costs (and minimal environmental impacts) are requirements for gasification systems proposed for utility applications. Operating costs can be minimized with a gasifier that is capable of operating on run-of-the-mine coals, including caking coals as well as fines. The gasifier should achieve high thermal efficiencies and high carbon utilization.

High efficiencies are obtained in the combined-cycle operation because efficient combustion turbines are combined with steam turbines. The gas turbine converts high-temperature (2300°F) heat efficiently, while the steam turbine utilizes low-temperature heat efficiently in the form of steam (1050°F).

Environmental impacts are minimized if the gasifier is designed and operated in such a

TABLE 17.12 Design and Actual Performance for the CW CGP*

Performance Parameter	*Design*†	*Actual Performance*		
			Illinois	*Pittsburgh*
Coal type	*SUFCO*	*SUFCO*	*No. 6*	*No. 8*
Sulfur content (wt%)	0.48	0.4	3.1	2.9
HHV (Btu/lb)	12,300	12,300	12,800	13,700
Coal-feed rate (TPD)	1000	1000	1000	1000
Oxygen consumption (TPD)	962	908	885	979
Gross power production (MWe)	114	116	120	125
Byproduct sulfur produced (TPD)	4.6	3.8	30	27
Coal/water slurry concentration (wt% solids)	60	60	64	62
Carbon conversion (%)	98.3	98.3	96.6	99.1
Gasifier/syngas cooler efficiency (%)	90.3	91.5	88.5	91.6
Gasifier refractory life (yr)	1.0	2.6	1.1	‡
Overall heat rate (Btu/kWh)§	11,500	11,300	12,000	11,600
Efficiency (%)	29.7	30.2	28.4	29.4

†Design parameters based on SUFCO coal.
‡Data not yet available.
§The heat rates were not optimized and are expected to be reduced by ~2500 Btu/kWh in future plants.
*"Coal Gasification: Direct Application and Synthesis of Chemicals and Fuels; A Research Needs Assessment," by Penner, S. S., and others for US DOE, Office of Energy Research; prepared under contract No. DE-AC01-85ER30076, NTIS-PR-360.[26]

way that tars are not formed. Low pollutant emission levels result from the combustion of a clean fuel. Pollutants are removed before combustion by using gas purification systems. Ash leaving the system should be in a benign form for disposal. On balance, coal gasification systems are environmentally superior to other alternative coal utilization technologies and will meet rigorous environmental standards for sulfur, nitrogen, and particulates.

A well-designed IGCC system can accept many types of feedstocks including lignites and petroleum coke. Gasification systems may be configured to produce other industrial chemical products, which adds flexibility. Table 17.12 presents design and actual performance data for the Cool Water Coal Gasification Process, which operated in the IGGC mode for five years.

Descriptions of Gasification Processes

General Electric Moving Bed. General Electric is operating a one-ton-per-hour, air-blown, advanced moving-bed gasifier in Schenectady, New York. The unit operates at 300 psig and is equipped with a mechanical stirrer to enable caking coals to be gasified. The fact that the volume of coal in the gasifier at any time is large compared to the steam and air injection rates allows considerable flexibility in handling feed interruptions. Figure 17.17 shows a diagram of the GE moving-bed gasifier.

Rotary-Kiln Gasifier. The rotary-kiln gasifier developed by the Allis-Chalmers Corporation consists of a tumbling bed of solids moving countercurrently to the flow of hot gases (Fig. 17.18). Reaction zones in the kiln parallel those in a moving-bed gasifier with the coal sequentially dried, devolatilized, gasified, and combusted. Steam and air are injected into the kiln through ports located beneath the tumbling solids. Although rotary kilns are commercially available, the gasifier is still in the development stage with a 600 ton/day demonstration plant being planned.

Westinghouse Fluidized-Bed Gasifier. Westinghouse has successfully operated a 15 ton/day air (or oxygen)-blown, fluidized-bed gasifier (Fig. 17.19). The untreated coal

Fig. 17.17. General Electric GEGAS-D advanced gasifier. (*Courtesy General Electric Co.*)

Fig. 17.18. Rotary ported kiln gasification process. (*Courtesy Allis-Chalmers Corp.*)

Fig. 17.19. Westinghouse pressurized fluidized-bed gasification system. (*Courtesy Westinghouse Electric Corp.*)

is fed to the gasifier where it is combusted in a stream of air or oxygen through the central feed tube. Steam fed with the air or oxygen and in the grid zone of the gasifier reacts with the coal and char to form hydrogen and carbon monoxide. As the bed of char circulates through the jet, the carbon in the char is consumed, leaving particles that are rich in ash.

Texaco Coal Gasification Process. The Texaco process (TCGP) offers the following design features: downward-loading, pressurized entrained-flow slagging reactor; air- or oxygen-blown process; high operating temperatures; flexible feedstocks and products; system coupling for cogeneration. The TCGP is designed to operate at pressures between 20 and 80 atm and temperatures between 1200 and 1500°C. High points in its development history are indicated in Table 17.13. Figure 17.20 shows a direct quench gas cooling configuration for the TCGP. The coal is wet-ground and mixed with water to make a slurry containing from 60 to 70 wt. % solids. The slurry is mixed with oxygen or air in the gasifier burner. By properly adjusting the oxygen/slurry ratio, temperatures are maintained above fluid ash temperatures.

The gases are cooled after leaving the gasifier, either by direct contact with quench water, as in Fig. 17.20, or by passage through a radiative cooler followed by a convection cooler. The former method is preferred when the output is NH_3 or H_2, as the required shift reactor is easily integrated in the quench mode. The cooled gases next go to a scrubber to remove char and ash. The product gases leaving the scrubber contain H_2, CO, CO_2, H_2O, and traces of Ar, N_2, CH_4, H_2S, and COS. Spencer and others have reported on the performance of the TCGP.[27,28]

Various products have been prepared in commercial plants, including H_2, NH_3, fuel gas for electricity generation, and synthesis gas (CO and H_2) for production of methanol and other oxo products.

Any carbonaceous material that can be formed into a concentrated, pumpable slurry can be gasified in this process. For example, petroleum coke is converted to synthesis gas in a 30-ton/day plant in Ube City, Japan,

TABLE 17.13 Development of the TCGP

Owner	*Type of Plant*	*Location*	*TPD*	*Start-up*	*Gas Cooling*	*Product*
Texaco, Inc.	Pilot	California	15	1973	d.q.[1]	—
Texaco, Inc.	Pilot	California	15	1978	d.q./h.r.[2]	—
Texaco, Inc.	Pilot	California	15	1981	d.q.	—
RAG/RCH	Demonstration	W. Germany	165	1978	h.r.	Oxo-chemicals
Dow Chemical	Demonstration	Louisiana	400	1979	—	Synthesis gas for electric power
TVA	Demonstration	Alabama	190	1982	d.q.	Ammonia
Tennessee-Eastman	Commercial	Tennessee	900	1983	d.q.	Acetic anhydride
Southern California Edison Co., etc. (Cool Water)	Commercial	California	1000	1984	h.r./d.q.	Synthesis gas for electric power
Ube	Commercial	Japan	1650	1984	d.q.	Ammonia
SAR	Commercial	W. Germany	800	1986	d.q./h.r.	Oxo-chemical/H_2
IHI	Pilot	Japan	15	1989	d.q./h.r.	—
Lntong	Pilot	China	15	1991	d.q.	—
LuNan	Demonstration/ commercial	China	360	1991	d.q.	Ammonia
Shogang	Commercial	China	1100	1993	h.r.	Steel products

*[1]d.q. = direct quench; [2]h.r. = heat recovery.

Fig. 17.20. Flow diagram for the Texaco Coal Gasification Process utilizing direct quench of gases leaving the gasifier. (*Courtesy Texaco.*)

which has been operating since 1982. In 1991, a mixture of coal and sewage sludge was successfully gasified.

KBW Process. In the KBW process (Fig. 17.21), powdered coal is entrained with steam and oxygen and, under slight pressure, is fed into the gasifier where gasification at about 2800°F. The slagged ash drops into a water seal, from which it is removed by a continuous scraper conveyor. The high temperature reportedly results in production of a gas mixture rich in the desirable constituents of CO and H_2, with no methane and only limited amounts of CO_2. Characteristically, coal yields a gas of 50 to 55 percent CO and 30

Fig. 17.21. Koppers Totzek gasifier. (*"Comparative Study of Coal Gasification Processes—Koppers-Totzek, Lurgi and Winkler," Koppers Co., Pittsburgh, PA.*)

to 35 percent H_2, on a dry basis. The exit gas is cooled to about 400°F in a waste heat boiler and cleaned to remove particulate matter.

COMBUSTION

In the United States, approximately three-quarters of the annual coal consumption is burned in boilers to generate steam for turboelectric plants in the electric utility industry. On a worldwide basis this figure is closer to one-half. However, the use of coal in small-scale combustion devices for direct heating is significant outside the United States. A summary of the most important direct combustion processes for coal is given in Table 17.14.

The use of coal in combustion processes has drawn increasing criticism from environmentalists because of concerns over its contribution to acid rain (from oxides of sulfur and nitrogen) and global warming (from its high CO_2 production per unit of energy), as well as associated emissions of particulates and trace elements.[31] Despite increasing environmental regulations, which account for about 30 percent of the cost of a new plant, coal is still the fuel of choice for new plants by electrical utilities in the United States.[32] In addition, the amount of coal used for combustion is likely to increase steadily over the next decade in developing countries such as China. The combustion of coal is also important in several eastern European economies, such as Poland. Consequently, there is a strong impetus for development of so-called clean coal technologies that are also highly efficient.

Combustion Technology

Remarkable advances have been made in the energy efficiency of electric utility plans because of improvements in the steam-generating equipment and turbines. A method of expressing the efficiency of the conversion of coal to electricity is the heat rate, which is a measure of the pounds of coal required to generate one kilowatt-hour of electricity. Of course, the heat rate depends on the heating value of the coal, and the numbers usually are based on an assumed value of 12,500 Btu/lb, which is common for a bituminous coal. On this basis the heat rate of a modern, well-maintained plant in 1990 was about 0.7 lb. This is compared to average values of 1.4 lb in 1937 and 0.8 lb in 1968. Some new designs project heat rates close to 0.6 lb of coal per kilowatt-hour, which corresponds to an overall thermal efficiency of 45 percent, based on coal having an as-fired heating value of 12,500 Btu/lb.[33]

Further improvements in thermal efficiency may be achieved with new systems of power generation now being developed. One such technology, known as MHD (magnetohydrodynamics), is expected to have a thermal efficiency of between 50 and 60 percent if coal is used as a fuel.[34] The principle of MHD is to burn the fuel under conditions that will give a combustion gas temperature of about 4900°F, add an easily ionizable salt such as potassium carbonate (seeding) so that the gas will have a relatively high concentration of free ions and electrons, and then pass this conducting gas through a magnetic field and draw off electrical energy through electrodes placed in the path of the gas. By analogy, the hot conducting gas moving through a magnetic field behaves like the armature in a conventional turbogenerator. A coal-fired MHD generator would discharge less sulfur and nitrogen oxides to the atmosphere than a conventional plant of equivalent size because the seeding salt would absorb these oxides, which would be recovered during regeneration of the seed. The higher thermal efficiency would also reduce the amount of CO_2 evolved per unit of energy produced.

However, there are still some major technical obstacles to overcome before a commercial MHD plant can be built.[33,34] One problem is the selection of materials to withstand the extreme temperatures in the gas and meet the requirements that two walls in the duct must be conductors, and the other two walls must be insulators. A second problem is recovery of the potassium used as the feeding material; some estimates indicate that economical

TABLE 17.14 Summary of Direct Combustion Coal Process (Smoot)*

Process Type	*Description*	*Extent of Coal Use In U.S. (% of Total Used)*	*Commercial Use*	*Scale Size (TPD)*[a]	*Coal Types*	*Coal Size*
Power station	Commercial electricity production	78–80%				
Pulverized	Rapid burning of finely grained goal		Common	1000–10,000	All	0.01–0.025 mm
Fluidized-bed[b]	Well-stirred combustion		Pilot plant	2000–8000	All	0.15–0.6 cm
Stoker	Mechanically fed fixed bed		Small	100	Noncaking	1–5 cm
MHD	Combustion energy capture by magnetic fields		Laboratory	800–4000		
Coal/oil mixture (COM)	Burning coal/oil mixtures in oil furnaces		Demonstration			
Industrial heat/steam	Industrial plant power	8–11%				
Pulverized			Small	1–100		
Fluidized-bed			Pilot	1–100	same as above	
Stoker			Common	1–100		
COM			Demonstration	1–100		
Domestic/commercial	Hand-stoked space heating	1%		0.005–0.05	Noncaking	3–10 cm
Transportation	Fuel for railroads	0.01–0.02%		0.1–1	Noncaking	

[a]TPD = tons/day.
[b]Very adaptable, exhibits high heat transfer in bed as well as low-level pollutant products.
**Source*: Excerpted from Smoot, L. D. in *Fossil Fuel Combustion*, William Bartok and Adel F. Sarafin (Eds.). John Wiley & Sons, New York, 1991. Copyright © 1991 by John Wiley & Sons, Inc. and used by permission of the copyright owner.

operation will require 99 percent recovery. The U.S. DOE supports a research program on MHD that includes tests on prototype MHD components and the use of MHD exhaust gases for steam generation.[32] This five-year program is intended to prove the feasibility of the concept by the end of 1993. Research on the MHD concept also is being actively pursued in other countries such as the Former Soviet Union, China, Japan, Poland, Italy, and India. The work in Russia has progressed to the pilot-plant scale.

A second promising technology is the combined-cycle plant, in which both a gas turbine and a steam turbine are used to generate electricity. In one version of this concept, the coal first is gasified to produce fuel gas, which is cleaned and then burned in a gas turbine. An example of such a plant is shown in Fig. 17.22. The hot exhaust gases subsequently are used to generate steam that drives a steam turbine. A second type of combined-cycle plant uses a pressurized fluidized-bed combustor (PFBC) to generate the hot gas for the gas turbine.[35,36] However, because of the relatively low maximum temperature limit on the PFBC (see below), the thermal efficiency is lower than for the gasifier-based plant. An important advantage is that the thermal efficiency of a combined-cycle plant exceeds 40 percent, whereas that of a conventional pulverized coal-fired boiler

Fig. 17.22. A combined-cycle plant in which the coal is first gasified prior to combustion. (Reprinted with permission from Schobert, H. H., *Coal: The Energy Source of the Past and the Future*, American Chemical Society, Washington, D.C., 1987. Copyright 1987, American Chemical Society.)

equipped with flue gas desulfurization is about 35 percent. The development of more advanced gas turbines, which allow higher gas inlet temperatures, will significantly increase the efficiency advantage of combined-cycle plants. A second aspect of these plants is the requirement for a clean gas system for the gas turbine. This is a technical challenge in terms of the development of processes to remove trace amounts of impurities from gasifier effluents without having to cool down the gas stream. However, this feature is also an advantage because the pollutants can be removed from a much lower volume of gas when compared to the stack and, in most cases, are in a form that is more easily removed. Another advantage of combined-cycle plants is their modular nature, which offers flexibility in construction and the ability to add capacity in smaller increments. The feasibility of the combined cycle plant was demonstrated at Southern California Edison's Cool Water station, which utilized a Texaco gasifier to generate the fuel gas.[27]

The use of coal slurries, either coal–oil mixtures (COM) or coal–water mixtures (CWM), has been the subject of considerable research because of its potential to replace fuel oil in boilers.[37,38] The recent work has shifted in favor of CWM over COM, as the cost savings for the latter were not enough to offset the conversion costs. CWM is typically composed of 60 to 75 percent coal, 24 to 39 percent water, and 1 percent chemical additives. The advantages of CWM compared to dry pulverized fuel are ease and safety of handling, transporting (e.g., via pipeline), and injection of the fuel. It is possible that the manufacture of CWM can be integrated into coal beneficiation processes such as froth flotation. The potential problems with the replacement of CWM for oil are: (1) limited flame stability and turndown; (2) incomplete combustion; and (3) slagging and fouling caused by inorganic constituents. This technology is still in the development stage, and work has slowed considerably since 1986 because of low oil prices. Several demonstration projects have been completed or are still under way.[38]

Combustion Equipment

Industrial and electric utility boilers are fired with either stokers, pulverized-coal-fired burners, or cyclone burners, the choice depending on the kind of coal and the amount of steam needed. Table 17.15 gives the kinds of coal that can be burned with the various kinds of firing. Although good results usually will be obtained with the respective coal and firing equipment, this table should be used only as a rough guideline, and final equipment selection should be based on a sound engineering analysis.

Stoker Firing. Stokers are the oldest continuous coal combustion devices, having been in operation for more than 150 years. They are still used to fire small-capacity boilers. Table 17.16 gives the approximate range of capacity for each kind of stoker. There are several types of stokers, which are classified by the method of coal feeding and the relative direction of the fuel and air. The three main

TABLE 17.15 Burning Equipment for Various Coals[99]

Fuel	*Stokers* *Under Feed*	*Traveling or Chain Gate*	*Spreader*	*Pulverized-Coal burner*	*Cyclone Burner*
Anthracite		X		X	
Bituminous:					
17–25% volatile	X		X	X	X
25–35% volatile				X	X
strongly coking	X	X	X	X	X
weakly coking		X	X	X	X
Lignite		X	X	X	X

TABLE 17.16 Approximate Range of Capacity of Stokers[99]

Type	*Steam, M lb/hr*	*Grate Heat Release, M Btu/hr sq. ft. (max.)*
Single retort	5–50	200
Multiple retort	40–300	300
Traveling or chain grate	10–300	300
Spreader	10–300	1000

types are underfeed, crossfeed, and overfeed. Both fuel and air have the same direction in retort stokers; this arrangement is called underfeed burning. The fuel moves across the air direction in chain- or traveling-grate stokers; this design is called crossfeed burning. The spreader stoker approximates overfeed burning, the incoming fuel moving toward the air. Except for certain types of coal gasifiers, in which lump coal moves downward toward a grate against air (or oxygen and steam) coming through the grate, no conventional combustion system operates purely in the overfeed mode.

An example of a crossfeed chain-grate stoker being used to heat a fire-tube boiler is given in Fig. 17.23. This stoker consists of an endless track of grate bars that pass over rotating sprockets at each end, similar to a conveyor belt. Coal is fed onto one end of the grate, using a hopper, to a depth of about 4 inches. The coal burns as it moves along the grate, and, at the far end of the grate, only the ash remains. The ignition of the coal is assisted by radiant heat from a refractory lining.

Pulverized-Coal Firing and Cyclone Firing. Electric utility and large industrial plants favor pulverized-coal-fired and cyclone-fired furnaces because of their inherent flexibility regarding the kind and quality of coal, their comparatively good availability, their quick response to load changes, and their extremely high steam-generating capacity. Recent units

Fig. 17.23. Example of a chain-grate stoker equipped with a five-tube boiler. (Reprinted with permission from Schobert, H. H., *Coal: The Energy Source of the Past and the Future*, American Chemical Society, Washington, D.C., 1987. Copyright 1987, American Chemical Society.)

Fig. 17.24. Methods of firing pulverized and crushed coal. (*Courtesy Bureau of Mines.*)

generate as much as 9.5 million lb of steam/hour at 3500 psig and 1000°F.

The burner and furnace configurations for the main type of pulverized-coal firing (often called suspension firing), and cyclone firing are shown in Fig. 17.24. There are some design variations in vertical, impact, and horizontal suspension firing, but these schematic drawings serve to illustrate the principles.

Coal-fired boilers are further differentiated by the operating mode: the "dry bottom" or "wet bottom" (slagging) mode. Most modern pulverized coal-fired boilers are of the dry bottom type, in which the ash is cooled before contacting any surface.[39] Wet bottom firing provides a more convenient method of ash disposal but has several disadvantages that have led to its virtual elimination in the United States.[35] The disadvantages include less flexibility in fuel selection, higher incidence of fouling and corrosion, and higher levels of NO_x formation.

The first suspension-fired furnace in the United States was designed like the one shown for vertical firing in Fig. 17.24. Pulverized coal (about 70% through a 200-mesh screen) is transported to the burner with primary air, the amount of this air being about 20 percent of that needed for complete combustion. The balance of the air, known as secondary air, is admitted through openings in the furnace wall. Because a large percentage of the total combustion air is withheld from the fuel stream until it projects well down into the furnace, the ignition stability is good.[40] This type of firing is well suited for coals that are difficult to ignite, such as those with less than 15 percent volatile matter. Although no longer used in central-station power plants, this design, with delayed admission of secondary air, may find favor again if low-volatile chars from various coal-conversion processes are burned for heat and power.

The other types of suspension firing use burners in which the primary air and coal and the secondary air are mixed just before or immediately after entering the furnace. With tangential firing, the burners are arranged in vertical banks at each corner of a square, or nearly square, furnace and directed toward an imaginary circle in the center of the furnace. This produces a vortex with its axis on the vertical centerline. The burners consist of an arrangement of slots, one above the other, admitting through alternate slots the primary air–fuel mixture and the secondary air. The burners can be tilted upward or downward 30 degrees from the horizontal plane, enabling the operator to control superheat and to permit selective utilization of furnace heat-absorbing surfaces. Basically, the turbulence needed for mixing the fuel and air is generated in the furnace instead of in the burners.

Cyclone firing is a form of two-stage wet bottom firing in which the coal is not pulverized as for suspension firing but is crushed to 4-mesh size and admitted tangentially with primary air to a water-cooled cylindrical chamber called a cyclone furnace. The finer particles burn in suspension, and the coarser ones are thrown by centrifugal force to the furnace wall. The wall, having a sticky coating of molten slag, retains the coal until it burns. The secondary air, which is admitted tangentially along the top of the furnace, completes the combustion of the coarse particles. Slag drains continuously into the main boiler furnace. The main advantage of cyclone firing is the ability to use larger coal

particles. However, being a wet bottom technique, it suffers from the disadvantages discussed above.

Fluidized-Bed Combustion. In the past 20 years, fluidized-bed combustion (FBC) has advanced from the laboratory and pilot-plant scale to commercial status worldwide.[36,39,41,42] Although this technology is not yet widely used in the electrical power industry, the potential advantages of this burning mode are clear: (a) The fuel-bed temperature is low, about 1550°F, which means less formation of nitrogen oxides, and retention of some of the sulfur in the ash of certain coals. (Adding dolomite or limestone to the fuel bed greatly improves sulfur retention.) It is equally important that there is less volatilization of sodium and potassium in the coal; consequently, there are fewer deposits on and corrosion of furnace, superheater, and reheater tubes. (b) Heat-transfer rates from the fluidized bed for immersed heat-transfer surfaces are relatively high, as much as 100 Btu/hour/ft^2/°F. Temperature control is easier, as the bed temperature is more uniform because of the rapid mixing in the bed. (c) The FBC has the capability to burn many kinds of coal and is less sensitive to changes in coal heating value or ash characteristics than other combustion methods. (d) The FBC is adaptable to several kinds of power generation cycles (e.g., cogeneration or combined cycle) can operate at atmosphere or elevated pressure, and can use bubbling or circulating beds with in-bed or external heat exchangers. This versatility allows engineers to develop FBC systems that are tailored to specific applications, which can result in significant cost savings when compared to other technologies.[39]

It is believed that the application of fluidized-bed boilers to electric-utility stations will reduce operating costs and increase efficiency compared with conventional pulverized-coal fired plants of the same capacity that are equipped with emission controls. The principle of the fluidized-bed combustor is illustrated schematically in Fig. 17.25.

Fig. 17.25. Basic components of a fluidized bed combustor. (*Shang, J. Y., Department of Energy, Morgantown WV.*)

The current trends in the area of FBC are as follows:[39] There is more interest in running units at elevated pressures (6–15 atm) in order to reduce equipment size, lower emissions, reduce sorbent requirements, and increase the thermal efficiency by the use of a gas turbine in addition to a steam turbine in the combined-cycle concept. There is also a trend away from bubbling beds and toward "circulating" beds, which use finer solids and higher velocities and are able to use sorbents more efficiently.[42]

Several large-scale demonstration projects of FBC technology for electrical power generation are now under way. These include a 110 MWe atmospheric pressure circulating fluidized-bed boiler at the Colorado–Ute Electric Association's Nucla Station, which began operation in August 1988. A second project is the American Electric Power Tidd 70 MWe pressurized fluidized-bed combustor/combined-cycle demonstration, which began generating power in December 1990. Both projects were partly funded by the Clean Coal Technology Program.[43]

Boiler Types

There are various kinds of industrial and utility boilers, broadly classified as fire-tube and water-tube. In the former, the hot combustion gases pass through tubes, and heat is transferred to water outside the tubes. The most common and least expensive boiler of this type is the horizontal return tubular (HRT) boiler. However, because of the design and construction of fire-tube boilers, there is a definite limitation to their size and the pressure that they can tolerate.

Water-tube boilers may be broadly classified at straight-tube and bent-tube types, the latter having several variations in design and being preferred for applications where higher capacities and steam pressures are required. In both types, heat is transferred by radiation or convection to the outside of the tubes, and water flows inside the tubes as a result of thermal circulation, or in the case of certain bent-tube boilers, as the result of forced circulation. A comparatively new version of the forced-circulation, bent-tube boiler for central station power plants is the "once through" type. The feed-water passes progressively through the heating, evaporation, and superheater sections; no drum is used for separating the steam from the boiler water as in other boilers, so the ratio of water circulated in the boiler to steam generated is unity. The data in Table 17.17 approximate the range of steam capacities and pressures for the principal types of boilers.

Pollution Controls

The emission of sulfur and nitrogen oxides and particulates from coal combustion has become a problem of increasing concern and regulation.[31,32] The combustion of coal contributes about 25 percent of the particulate matter, 25 percent of the sulfur oxides, and 5 percent of the nitrogen oxides in the atmosphere.[33] A major concern about the particulate matter, most of which is inorganic matter derived from the mineral content of

TABLE 17.17 Approximate Range of Capacities of Various Types of Industrial and Utility Boilers[99]

Type	*Capacity, lb Steam/hour*	*Maximum Design Pressure, psig*
Fire tube (HRT)	1000–15,000	250
Water tube		
straight	15,000–150,000	2000
bent, 3-drum, low-head	1000–35,000	400
2-drum, vertical	1000–350,000	1000
electric utility	up to 9,000,000	3500

the coal, is that the smallest particles are respirable and may pose a health hazard. The collection of particulate matter is done in most power plants by the use of electrostatic precipitators, which have been developed to very high efficiencies (>99%). Other techniques for particulate removal include baghouses and cyclone collectors.

The significant reduction of SO_2 and NO_x emissions from coal combustion by the year 2000 is a major part of the 1990 Clean Air Act. The achievement of these goals will be assisted by a U.S. DOE program known as the Clean Coal Technology Program, which provides government funding for a range of advanced pre-combustion and post-combustion technologies.[43] The current technology for removal of SO_2 is Flue-Gas Desulfurization (FGD), in units known as scrubbers. The most common scrubbers use a slurry of lime or limestone to capture the sulfur oxides, and the product is a sludge containing calcium sulfite and calcium sulfate. However, the disposal of sludge is another environmental problem. A more favorable solution for newer plants is to use fluidized-bed combustion with a sulfur capture agent or a combined-cycle system where the sulfur can be more easily removed from the gasification stream.

The major component of nitrogen oxides emitted from coal is nitrogen oxide (NO). Most of the nitrogen formed from combustion of coal does not originate from the coal but from the reaction of nitrogen and oxygen in the air that is used to burn the coal. Consequently, the control of nitrogen oxides can be most easily achieved by changing the conditions of the combustion process. This usually involves the use of staged combustion where the coal first is burned in a fuel-rich flame that subsequently is made fuel-lean. This is made a viable option for an existing plant by changing the burner design or by the addition of natural gas.

Several techniques are being developed for post-combustion scrubbing of NO_x, such as selective catalytic reduction. This technique now is used in Japan and Germany but is not yet widely used in the United States because of its high cost.[31] The use of a fluidized-bed combustor where the combustion is carried out at lower temperature in the presence of a sorbent will reduce both nitrogen oxides and sulfur oxides.

DIRECT COAL LIQUEFACTION

The primary purpose of any coal liquefaction process is to produce a liquid product with an increased hydrogen-to-carbon ratio, a reduced heteroatom (e.g., S, O, N) content, and a minimal inorganic content. This has been accomplished by a variety of methods, which can be divided into the three categories of pyrolysis, indirect liquefaction, and direct liquefaction.[96]

Direct coal liquefaction (DCL) refers to the process of converting coal to liquid products by mixing the coal with a recycle process solvent or petroleum-derived residual oil and reacting it under hydrogen at 750 to 850°F (400–450°C) and 1000 to 2500 psig (7–17 MPa). The process involves the breakdown of the coal structure into soluble products promoted by hydrogenation. The solubilized products, consisting mainly of aromatic compounds, then may be upgraded by conventional petroleum refining techniques. For information on the early history of DCL see Wu and Storch,[44] Donath and Hoering,[45] and Ministry of Fuel and Power.[46]

The success of any DCL method is highly dependent on the type of coal used in the process. Coal rank has been shown to influence both the overall conversion and the product distribution. Lower-rank coals have been reported to give birth higher and lower conversions than bituminous coals,[47] with most evidence supporting the latter. However, the liquid products from subbituminous coals are generally more volatile and of lower molecular weight[48] than those for higher-rank bituminous coals.

After the end of World War II, the United States adopted much of the technology developed by Germany in the construction of a 200 to 300 bbl/day plant at Louisiana, Missouri, operated by the Bureau of Mines

from 1949 to 1954.[44] Although the Germans had demonstrated the commercial-scale operation of DCL plants, coal-derived liquids could be produced only at high cost. Research in the United States was limited prior to the 1970s, primarily because of the availbility of vast reserves of inexpensive oil discovered in the Middle East in the late 1940s. However, the 1973 oil embargo created a surge in interest in coal liquefaction as one of a number of routes to produce petroleum substitutes from indigenous resources. From this renewed effort in the early 1970s, four major processes evolved: the H-Coal process, the Exxon Donor Solvent (EDS) process, and two Solvent Refined Coal (SRC-I and SRC-II) processes. All were single-stage processes that attempted to convert coal to liquids in a single reaction stage. By the late 1970s and early 1980s these processes were being developed on the pilot-plant scale. This section discusses these processes as well as two-stage processes and future research in DCL.

The H-Coal Process

The H-Coal process was an adaptation of the H-Oil process used in the petroleum industry to convert heavy oil residues to lighter fractions using a catalytic ebullated bed reactor.[49] Research on the H-Coal process began in 1964 at Hydrocarbon Research, Inc. (HRI) on a bench-scale unit, and by 1973 the design of a 200- to 600-ton/day pilot plant was under way. Construction of the plant was completed in early 1980. Located adjacent to the Ashland Oil refinery in Cattletsburg, Kentucky, it operated succesfully from startup in May 1980 until shutdown in November 1982.

H-Coal employed a single catalytic stage to produce a synthetic crude oil.[50,51] Figure 17.26 shows a schematic of the H-Coal process. The coal is crushed, dried, and mixed with a recycle oil and hydrogen gas before being preheated to approximately 850°F (454°C). The preheater effluent then is fed to the bottom of the ebullated bed reactor. During operation, fresh catalyst (an extruded cobalt–molybdenum catalyst) is fed to the top of the reactor while spent catalyst is removed from the bottom of the reactor to maintain constant inventory. The constant upward flow of the coal slurry causes the catalyst to remain in a fluidized state. The products are removed from the top of the reactor, and the various fractions are separated by conventional techniques.

The H-Coal products had a high H/C ratio and a low heteroatom content, as shown in Table 17.18. The process design allowed for a high throughput rate. However, the use of high temperatures in the reactor caused excessive thermal cracking, resulting in high gas yields and, correspondingly, low hydrogen utilization efficiency. Typically, between 14,000 and 20,000 standard cubic feet of hydrogen were required for each ton of coal processed.

The H-Coal process produced approximately 3 to 3.5 barrels of liquid product for each ton of coal feed.[49] Testing showed that the process was best suited for high-volatile bituminous coal feeds; the use of low-rank coals reduced throughput and distillate yields considerably. The demonstrated success of the ebullated bed reactor in this system led to its future use in two-stage systems.

The Exxon Donor Solvent Process

Research on the Exxon Donor Solvent (EDS) process began at Exxon in 1966 and by 1975 had progressed from investigations using 100 cm^3 batch units to a continuous $\frac{1}{2}$-ton/day pilot plant.[52] With the help of external funding, a 250-ton/day pilot plant was designed and constructed. The Exxon Coal Liquefaction Plant (ECLP) at Baytown, Texas began operations in April 1980, and was shut down and dismantled in late 1982.[51]

A schematic of the EDS process is shown in Fig. 17.27.[50,51] The coal is ground and mixed with a hydrogenated distillate recycle solvent before entering the liquefaction reactor. Hydrogen is added to the slurry in a tubular thermal reactor, which operated at 800 to 880°F (425–470°C) and 1500 to 2000 psig (10–14 MPa). The total reactor product is set to a distillation unit, where it is separated into gas, naphtha, distillates, and a vacuum

Fig. 17.26. Schematic of H-Coal process. (*Courtesy The Pace Company, Denver, CO.*)

TABLE 17.18 Heteroatom Content of Product Streams from Various Direct Liquefaction Processes

Process	*Wt. % in Naphtha (C_5—390°F)*				*Wt. % in Distillate (400–975°F)*			
	O	*N*	*S*	*H*	*O*	*N*	*S*	*H*
H-Coal	0.6	0.2	0.07	13.79	1.1	0.5	0.08	9.52
EDS	2.42	0.23	0.28	12.03	2.0	0.6	0.3	8.97
SRC-II	3.51	0.42	0.26	10.46	3.63	0.99	0.25	9.27
ITSL	0.41	0.14	0.015	13.78	0.56	0.20	0.015	10.21

bottoms slurry. The 400 to 850°F (204–455°C) distillate fraction is sent to a fixed-bed catalytic solvent hydrogenation unit to generate a hydrogenated H-donor recycle solvent, which supplies hydrogen to the coal in the liquefaction reactor. The use of an external reactor to hydrogenate the recycle solvent distinguished the EDS process from other single-stage processes.[53] Additional hydrogen is supplied from the gas phase. The production of distillates is achieved by thermal cracking and hydrogenation in the liquefaction reactor. Although this process used a separate solvent hydrogenation unit, it is considered a single-stage process because both dissolution of the coal and upgrading of the dissolution products took place in the same reactor.

The operating conditions for the EDS process were similar to those of the H-Coal process, and, similarly, it was found that the process worked best with bituminous coals. The product distribution was slightly worse than obtained in the H-Coal process, and the distillate yield was lower, around 2.5 to 3 barrels of liquid product per ton of feed. The rather inferior performance was probably attributable to the absence of a catalyst in the liquefaction step of the EDS process. However, because the reactor was less complicated for the EDS process and the solids removal was simpler, the economics of the processes were roughly equivalent.[50]

The EDS process suggested that coal could be adequately hydrogenated in the absence of added catalyst. However, a later modification used vacuum bottoms recycle, in addition to the normal distillate solvent recycle, and displayed improved operating characteristics and higher distillate yield, particularly for low-rank coals. The increased residence time

Fig. 17.27. Schematic of Exxon Donor Solvent process. (*Courtesy The Pace Company, Denver, CO.*)

of the resid in the reactor increased the yield of naphtha range materials.[54] In addition, the ash in the bottoms recycle may have contributed some catalytic effect, resulting in increased conversion for lignite and subbituminous coals.[55] The recognition of these factors also was important to the development of the later two-stage liquefaction concept.

The Solvent Refined Coal Processes

The solvent refined coal (SRC) process began as a study on the feasibility of coal de-ashing by Spencer Chemical in 1962.[50] The aim was to produce a cleaned combustible fuel that was solid at room temperature. A 50-ton/day plant, located in Fort Lewis, Washington, operated in the SRC mode from its completion in October 1974 until it was modified in late 1976. It later was shut down in 1981. With this change came the new designation SRC-II, and the original process was sometimes designated as SRC-I. In 1972, another 6-ton/day pilot plant was built, in Wilsonville, Alabama, by Catalytic Inc. for EPRI and Southern Company Services Inc. This plant began operation in January 1974, and it is the only DCL plant currently in operation in the United States, although its present form is a highly modified configuration of the original.

Figure 17.28 shows a schematic of the SRC-I process. The feed coal is crushed and mixed with a recycle solvent and hydrogen prior to entering the preheater. The preheater effluent, at 700 to 750°F (370–400°C), then is fed to the dissolver unit, or thermal liquefaction unit (TLU), which operates at 840 to 870°F (450–465°C). There is no upgrading step, as the desired product is a solid at room temperature and not a distillate. The solids removal from the liquid slurry is accomplished by critical solvent de-ashing (CSD). The solids-free resid from the CSD was separated by vacuum distillation into a recycle solvent (the light fraction) and a solvent refined coal product (the bottoms).

The primary product of the SRC-I process was a low-sulfur, low-ash, solid boiler fuel

Fig. 17.28. SRC-I process for solvent refined coal. (*Courtesy The Pace Company, Denver, CO.*)

with a melting point of approximately 300°F (150°C), a Btu content of 16,000 Btu/lb, and an H/C ratio approximately the same as that of coal. Although this was considered to be a thermal liquefaction process, it was most successful using bituminous coals with high native pyrite content. Its beneficial effects certainly can be attributed to the now well-known catalytic properties of iron sulfides.

The SRC-I process suffered from a number of problems. Among them was the fact that the distillate solvent was a poor hydrogen transfer medium, and components of the solvent were incorporated into the resid, making it virtually impossible to maintain solvent balance.[51] Problems with hydrogen transfer often led to problems with retrograde reactions (i.e., coking), particularly in the preheater section. Attempts to correct this problem by moderating the conditions led to excessive quantities of toluene insolubles precipitating downstream in the CSD.

The SRC-II process was again a thermal process, relying only on the native mineral matter in the coal as a catalyst. However, in SRC-II part of the product from the dissolver was included with the recycle solvent, without first de-ashing. This helped maintain a high mineral matter concentration in the reactor, which must have augmented the extent of catalytically promoted reactions. The other changes in the SRC-II process can be seen in Fig. 17.29.

The SRC-II process operated under more severe conditions than SRC-I and gave primarily a distillable product, although gas yields also were high. This poor selectivity resulted in excessive hydrogen consumption. More than the other single-stage liquefaction processes, SRC-II was limited to the use of bituminous coals with high catalytic mineral matter content. Processing of subbituminous coals was possible only with the use of additional pyrite catalyst added to the feed stream.[48]

Two-Stage Processes

During the late 1970s and early 1980s, there was an increasing awareness that the process of direct coal liquefaction proceeded in two loosely defined steps, each requiring different reaction regimes. The first step involves coal dissolution, in which the coal is converted to a soluble form with high molecular weight and relatively little change in average composition from that of the coal. The second involves upgrading, in which the dissolved

Fig. 17.29. SRC-II process for solvent refined coal. (*Courtesy The Pace Company, Denver, CO.*)

products are converted to lower-boiling liquids with reduced heteroatom content.[56] Studies have shown that the conversion to tetrahydrofuran (THF) solubles can occur in one to five minutes, and that further reaction generally leads to unwanted hydrocarbon gas production.[50] It also was found that in the dissolution stage most of the hydrogen is supplied by the solvent rather than the gas. Therefore, the use of a short contact time (SCT) thermal liquefaction reactor followed by a catalytic upgrading reactor should optimize conditions for both reactions, leading to higher yield, more efficient use of hydrogen, and a better product distribution. These studies were the beginning of the two-stage liquefaction process.

The first process that attempted to separate the two steps was the nonintegrated two-stage liquefaction (NTSL) process.[57] This involved the use of a second-stage upgrading reactor placed downstream of an SRC-I dissolution reactor. Because of the high ash content of the dissolution product, which could cause plugging and rapid deactivation of the catalyst, a fixed-bed catalytic upgrading reactor could not be used.[58] The use of a variation of the ebullated bed type reactor, called the LC-Finer, installed at the Wilsonville facility, allowed the upgrading of a portion of the SRC-I resid, but the process still was considered inefficient, primarily because of the poor hydrogen transfer ability of the distillate solvent. Distillate yields were only about 40 percent (maf coal).[59] Use of the term "nonintegrated" stems from the fact that the hydrotreater does not contribute solvent to the first-stage reactor, as shown in Fig. 17.30.

Lummus Integrated Two-Stage Liquefaction Process

Several companies began work on integrating the two stages of the liquefaction process by recycling part of the hydrogenated resid. Around 1975, the Lummus Company began work on one such process, called integrated two-stage liquefaction (ITSL).[58,60] The process, shown in Fig. 17.31, consists of an SCT thermal liquefaction reactor followed by an LC-Finer, similar to the NTSL process. Aside from the recycle, the unique feature of this process configuration is the use of the Lummus antisolvent de-ashing (ASDA) unit, originally located between the SCT reactor and the LC-Finer, to remove ash from the resid prior to upgrading. The liquid product of this process is a below-650°F (343°C)

Fig. 17.30. Schematic of nonintegrated two-stage liquefaction process.

Fig. 17.31. Schematic of Lummus ITSL process.

distillate, with the above-650°F (343°C) residue being recycled as solvent.

The SCT reactor is essentially equivalent to the preheater in the SRC-I process. The coal slurry, consisting primarily of liquid products, solids (ash and unconverted coal), and gas, exits the reactor after a residence time of 2 to 12 minutes, depending on the feed coal rank, at 800 to 840°F (425–450°C) and 1500 to 2500 psig (10–17 MPa). The short contact time resulted in lower hydrogen consumption, similar conversion, and a resid that was more reactive to upgrading at milder conditions than SRC-I.[61]

The ASDA unit removed the ash and heavy components of the resid by agglomeration with process-derived naphtha. The process required the solids content of the reject stream to be less than 55 percent in order for it to be pumpable.[50] As mentioned, the ASDA originally was placed between the SCT reactor and the LC-Finer. Subsequent studies showed that, by placing the ASDA after the upgrading stage, the conversion yields were increased.[60] In studies with subbituminous coal, additional improvements in yield were achieved by recycling part of the ASDA feed to the SCT reactor. Relevant to future process development was the finding that the catalyst in the upgrading stage was not deactivated by the ash in the SCT reactor effluent.

The LC-Finer, also known as the hydrotreater (HTR), was used to upgrade the resid from the SCT reactor and to produce a hydrogenated recycle solvent. The low temperature of 700 to 750°F (370–400°C) and pressure of 2500 to 2700 psig (17–19 MPa) in the HTR prevented excessive gas production, but still allowed adequate hydrogenation to take place. The catalyst also was less prone to deactivation by coking at these low temperatures. Distillate yields of 52 percent (maf coal) were achieved by this process on a PDU scale.[58] The liquid product produced by the ITSL process had a better quality, as shown in Table 17.19, and was more easily upgraded by conventional petroleum reforming techniques than the single-stage process products.

The Kerr-McGee Company developed a similar ITSL process.[50,60] The conditions in the SCT reactor were more severe than in the Lummus ITSL, resulting in significant distillate production in the SCT reactor. Bench-scale studies showed that the high temperature in the SCT reactor also increased gas production.

Although the first stage of this system was considered a thermal reactor, it was found that the addition of 2 wt. % Fe_2O_3, with dimethyldisulfide (DMDS), to a bituminous coal feed resulted in a significantly increased yield. This is attributed to catalysis by the iron sulfides formed during the reaction.

Wilsonville Integrated Two-Stage Liquefaction Process

By the mid-1980s, the configuration of the 6-ton/day plant at Wilsonville, Alabama, had been significantly modified from the original SRC-I mode. The Kerr-McGee company supplied a CSD to replace the filtration units previously in use. In 1981, HRI, Inc. installed

TABLE 17.19 Distillate Yields from Various Direct Liquefaction Processes

Process	*Configuration*	*Distillate wt. % maf coal*	*Yield, bbl/ton maf coal*	*Distillate Quality, gravity °API*	*Heteroatoms wt. %*		
					N	*O*	*S*
SRC-II (1982)	one-stage noncatalytic	41	2.4	12.3	1.0	2.33	0.33
H-Coal (1982)	one-stage noncatalytic	52	3.3	20.2	0.50	1.0	0.20
Wilsonville (1985)	ITSL thermal-catalytic	62	3.8	20.2	0.25	1.9	0.23
Wilsonville (1986)	CC-ITSL catalytic-catalytic	70	4.5	26.8	0.16	<1	0.11

an H-Oil ebullated bed reactor. This allowed, for the first time, the testing of the TSL concepts on a large scale. The plant piping was changed to test the Lummus ITSL process configuration.[50,57,60] The TLU was bypassed, leaving just the preheater, and the de-asher was installed before the second stage. This configuration resulted in a total conversion of 88 percent (maf coal) with a distillate yield of 54 percent (maf coal). To reduce the rejection of organic material with the ash in the CSD, the TLU was reinstalled. This meant that most of the conversion would take place in the first stage, with the second stage serving as a solvent hydrogenation unit, resulting in easier solid–liquid separation. With this change, a distillate yield of 57 percent (maf coal) was reported with an Illinois #6 bituminous coal.[62]

Reconfigured Integrated Two-Stage Liquefaction Process

Research at Lummus showed that moving the ASDA unit downstream of the HTR increased the distillate yield.[50] The Wilsonville plant was modified to operate in the reconfigured integrated two-stage liquefaction (RITSL) mode by moving the CSD unit downstream of the second stage, as shown in Fig. 17.32.[57] In this configuration, the de-ashed resid and some hydrotreated distillate were combined as recycle solvent. As indicated by the Lummus research, no detrimental effects to the catalyst in the HTR were seen due to the ash-containing feed. The process originally achieved a distillate yield of 62 percent (maf coal). A further increase in distillate yield to 63 percent (maf coal) was accomplished by using an ash-containing recycle. The primary gain from this configuration was the reduction of CSD feed by 50 percent, thereby reducing the load on the unit and the amount of organics rejected with the ash.

Close-Coupled Integrated Two-Stage Liquefaction Process

The purpose of the close-coupled integrated two-stage liquefaction (CC-ITSL) process was to minimize the residence time between reactors, thereby improving the thermal efficiency, reducing the potential for retrograde reactions, and producing a higher-quality solvent.[50,62,63] Also, because the CSD had been moved downstream of the HTR, there was no need for pressure letdown followed by repressurization between the first and the second stages.[57] The CC-ITSL configuration, shown in Fig. 17.33, demonstrated that the inclusion of the gases and light oils from the TLU in the HTR feed did not adversely affect the process. Distillate yields from the CC-ITSL of up to 64 percent (maf coal) were reported for Illinois #6 bituminous coal.[62]

Fig. 17.32. Schematic of reconfigured ITSL process.

Fig. 17.33. Schematic of close-coupled ITSL process.

Catalytic Two-Stage Liquefaction Process

HRI began research on a catalytic two-stage liquefaction (CTSL) process in 1982.[50] In this process the first and second stages were close-coupled catalytic ebullated bed reactors, originally using supported $CoMo/Al_2O_3$ catalysts.[60] The temperature of the first stage was lowered to 750°F (400°C) to permit a closer balance between hydrogenation and cracking rates. This also allowed in situ solvent hydrogenation. In contrast to the ITSL, the temperature of the CTSL second stage was higher (around 825°F, 440°C), than that of the first stage. This was intended to promote hydrocracking of the resid in the HTR, and to produce an aromatic recycle solvent. Later, the $CoMo/Al_2O_3$ catalyst was replaced in the first stage with a $NiMo/Al_2O_3$ catalyst, while the second stage continued to use the cobalt-molybdenum catalyst. The product from the second stage had effectively been catalytically hydrogenated twice and was much lighter than the product from previous processes. Therefore, pressure filtration was used for solids removal, with a higher efficiency for solid–liquid separation than could be achieved by CSD. The distillate yield for the CTSL process was as high as 68 percent (maf coal) for an Illinois #6 bituminous coal.[60]

The Amoco Oil Company developed an ITSL process very similar to HRI's CTSL process.[60] The main differences were the use of a higher-temperature, smaller-volume, thermal first-stage TLU and a lower temperature in the catalytic hydrotreater. This process was tested on HRI's bench-scale unit. The higher temperature in the TLU (820–850°F, 438–455°C) resulted in a 30 percent higher gas yield than HRI's CTSL process, whereas the lower temperature in the HTR (ca. 700°F, 370°C) contributed to longer catalyst life. However, the distillate yield for the Amoco process was considerably lower than that of HRI's CTSL process.

In 1985, the Wilsonville plant installed an additional ebullated bed reactor to replace the existing TLU.[50] This allowed the plant to operate without (Thermal/Catalytic, T/C CC-ITSL) or with (Catalytic/Catalytic, C/C CC-ITSL) a first-stage catalyst.[57] It is important to note that in the T/C CC-ITSL runs with subbituminous coal, iron oxide and dimethyldisulfide were added to the feed slurry as disposable catalysts, although the operational mode is still termed "thermal." This addition also was used for some bituminous coal runs, resulting in improved overall conversion and distillate yield.[57] The Wilsonville C/C CC-ITSL is slightly different from the HRI CTSL in two ways: the temperature in the first stage reactor is 800 to 820°F (425–438°C), and the second stage is 760 to 795°F (404–424°C). This allows solvent hydrogenation in the second stage rather than the first stage, similar to the ITSL process.

The second deviation from the HRI CTSL is the continued use of the CSD unit to remove the ash and solids. Improvements in the CSD unit have reduced organic rejection to the same levels that HRI achieves with filtration. These changes have allowed the Wilsonville plant to achieve distillate yields of up to 77 percent (maf coal) with a cleaned Ohio #6 bituminous coal.[64] More typical yields of 70 percent (maf coal) were found with Illinois #6 bituminous coal.[64] Recent studies demonstrated the beneficial effects on the distillate yield of deep-cleaning the coal, achieving a distillate yield of 74 percent with a Pittsburgh #8 bituminous coal.[65]

Research in DCL is not limited to the United States. Currently the British Coal Corporation (BCC) has constructed a 2.5-ton/day PDU at Point of Ayr, North Wales.[66,67] The facility, using a two-stage liquefaction process called Liquid Solvent Extraction (LSE), is supported by the United Kingdom Department of Energy, European Economic Community, Ruhrkohle, and Amoco. In the LSE process, dried crushed coal is slurried with a high-grade recycle solvent, pressurized to 300 psig (2 MPa), heated to 770°F (410°), and fed to the thermal first-stage reactor. The low pressure in the first stage is possible because of the excellent hydrogen donor properties of the recycle solvent. Solids are removed from the dissolved coal stream by filtration, producing a low-ash

solids-free filtrate called the coal extract solution. The coal extract solution is pressurized to 3090 psig (21 MPa), mixed with hydrogen gas, and fed to the catalytic, ebullated-bed, second stage, where it is hydrocracked at 750 to 850°F (400–450°C).

DCL and the Future

Like that of the past, future research primarily will be driven by economic considerations. The use of large-scale commercial DCL will be feasible only when the costs of coal-derived products are comparable to those of petroleum-derived products. Special circumstances, such as those present in Germany during World War II, may occasionally make DCL economically feasible, but these situations are usually temporary. Two recent studies[50,60] reported that, at an equivalent crude petroleum cost of around $35/bbl, coal liquefaction would be economically advantageous. However, with the current (1992) cost of crude oil closer to $20/bbl, the viability of a large-scale DCL operation is tenuous, to say the least. Current research is focused on improving the economics of DCL. Several approaches to reducing the cost are currently under investigation.

Reducing the cost associated with hydrogen production will greatly improve the economics of any coal liquefaction scheme. This may be done either by improving the hydrogen-generation methods or by improving the efficiency of hydrogen use (e.g., minimizing hydrocarbon gas make). Minimizing the severity of the reaction conditions, while maintaining a high yield, will reduce both equipment costs and gas production, thereby reducing hydrogen consumption.

Coal beneficiation, which involves the separation of some of the ash from the coal prior to liquefaction, may yield several benefits. Primarily it will reduce the load on the solid–liquid separator, thereby reducing the amount of organics rejected with the ash. Another benefit will be in the reduction of wear on plant components and better utilization of reactor volume. However, removal of mineral matter prior to reaction will eliminate any beneficial catalytic effect that the mineral constituents may contribute.

Improved catalysts, in both stages of a two-stage liquefaction system, probably represent the area of greatest potential benefit to the process economics.[68] Coal conversion reactions are generally slow, and any improvement in reaction kinetics could result in a large economic return. Currently research is being devoted to developing a dispersed slurry phase, as well as supported, catalysts that are more active and selective, and resistant to deactivation in the presence of high-boiling feeds.

The use of less expensive western U.S. subbituminous coals as feedstocks could further improve the process economics. It has been shown that they can produce lower-boiling and higher-quality liquids,[44,48,55] and that they cause less rapid deactivation of the second-stage catalyst.[60] The main problem is their lower rate and extent of conversion, which causes inefficient use of reactor space and the loss of additional product on solids separation. The different behavior of low-rank coals is believed to relate to their thermal sensitivity and high propensity for cross-linking. Dissolution promoted by an effective slurry-phase catalyst may be able to counter this tendency through increasing the effectiveness of hydrogenation.

The solution to the problems of making DCL economically viable may exist as one, or any combination, of the above techniques. There is also the possibility of discovering approaches that represent a radical departure from conventional techniques. Meanwhile, research is continuing to move toward the ultimate goal of an inexpensive alternative to imported petroleum.

PETROCHEMICAL FEEDSTOCKS

Feedstocks for the petrochemical industries now are produced mainly from crude oil (Chapter 15) and natural gas (Chapter 16). About 90 percent of all organic chemicals are manufactured from only six feedstocks: synthesis gas, ethylene, propylene, butadiene, benzene, and *p*-xylene, with synthesis gas

accounting for 56 percent of these feedstocks. Synthesis gas (a designation for various mixtures of carbon monoxide and hydrogen) now is produced by steam reforming of natural gas or light oils and is used in the manufacture of ammonia, methanol, and other chemicals. However, synthesis gas also may be produced by the gasification of coal.

There are many references that discuss and evaluate processes for the production of chemicals from coal.[7,90,95,98] A typical flow sheet for a plant is shown in Fig. 17.34. This "Coal Refinery" used hydroliquefaction, pyrolysis, and gasification and produces a product slate as shown in Fig. 17.35. The acronym POGO refers to Power Oil Gas Other, a project of the U.S. Department of Energy.

Most basic chemicals currently produced from oil can be produced from coal using demonstrated technology. The broadest range of synthesized products is available by conversion of coal to synthesis gas and utilization of existing processes for production of chemicals such as ammonia, methanol, and acetyl derivatives. Advances in the chemistry of synthesis gas are rapidly expanding the array of chemicals obtainable from coal by this route.[69-71] Many other chemicals can be separated from the by-products of gasification and pyrolysis. In fact, the foundation of the modern chemical industry was built on processes utilizing coal hydrogenation and pyrolysis to obtain a variety of chemicals including dyes, solvents, fuels, and pharmaceuticals.[72] However, in the 1950s the wide availability and low cost of petroleum pushed coal into the background as a raw material.

Today, the choice of feedstock for chemicals production depends on complex technical, economic, environmental, and political factors. Clearly, not all chemical products are suitable for production from coal with current technology. Some factors that must be considered in the evaluation of the appropriate feedstock for a particular chemical product are: (1) the correspondence between the carbon/hydrogen ratio in the chemical and the feedstock, (2) the delivered cost of alternative raw materials, (3) capital costs, (4) environmental protection, and (5) the reliability of supply. Except for special situations, the manufacture of chemicals from coal at present coal prices relative to the prices of petroleum and natural gas does not provide sufficiently attractive economic benefits. Special situations typically are a result of the concurrence of several favorable factors, such as the location of manufacturing facilities very close to coal mines with coal types that closely match the needs of the desired operation, coupled with the use of a process that can best utilize the composition of such coal with minimal conversion toward the hydrogen-to-carbon ratio of competing feedstocks. Alternatively, other situations arise because of political considerations where government subsidies or a closed economy provide the incentives for development of a synthetic fuels or a coal-based chemicals industry.

Chemicals from Coal

The use of coal as a raw material started with the discovery of the first coal tar color by Sir William Perkins in 1856. Prior to this time the liquid by-products of coal coking were regarded as wastes. Perkins's discovery revealed that valuable materials could be isolated or made from coal tar and ultimately led to the establishment of the modern chemical industry. Coke ovens rapidly came to be important as sources of chemicals such as benzene, naphthatlene, ammonia, and hydrogen. Coal tar became an important source for aromatic and heterocyclic chemicals for the dye and pharmaceutical industries.

Chemical by-products from coke manufacture amount to roughly 4 percent of the coal feed. Although this yield initially supplied adequate amounts of chemicals for developing end uses, the growth of the industry eventually required additional sources of chemicals. The advent of technology for producing calcium carbide and acetylene in 1892 provided a means for conversion of coke to chemical products.[73] Also, this technology stimulated the growth of an independent chemical industry by allowing the manufacture of chemicals to be pursued apart from the manufacture of coke for steel production.

Fig. 17.34. Simplified block diagram of the POGO complex. (*O'Hara, J. B., Klumpe, H. W., Bela, A., and Jentz, N. E., "Project POGO—A Coal Refinery,"* Chem. Engr. Prog., **74**, 8, *pp. 49–63, Aug. 1978. By permission.*)

Fig. 17.35. Projected POGO plant yields. (*O'Hara, J. B., Klumpe, H. W., Bela, A., and Jentz, N. E., "Project POGO—A Coal Refinery,"* Chem. Engr. Prog., **74**, 8, *pp. 49–63, Aug. 1978. By permission.*)

In Germany, shortages of petroleum during World Wars I and II prompted industry to develop processes for the conversion of coal into liquids for use as motor fuels. The two most important processes developed and implemented to meet wartime requirements were the Bergius coal liquefaction process and the Fischer-Tropsch process.[74] The Bergius process is a direct liquefaction process involving high-pressure catalytic hydrogenation. The Fischer-Tropsch process is an indirect liquefaction method in which synthesis gas reacts over a metal catalyst to form a range of hydrocarbon products including oxygenates and olefins. Because of the reactive nature of the olefinic and oxygenated coproducts, the Fischer-Tropsch process is particularly appropriate for coproduction of fuels and chemical intermediates. The history of chemicals production from coal shows that coal chemicals typically have been obtained as by-products from processes primarily designed to produce such products as motor fuel, town gas, or coke. Except under certain favorable conditions, economics still suggest that the production of chemicals must be coupled with the manufacture of other end products.

Coke production remains the most important demonstrated technology associated with the direct production of chemicals from coal. Industrial chemicals currently obtained in significant amounts from this source include benzene, toluene, xylene, naphthalene, anthracene, phenanthrene, phenol, ammonia, ammonium sulfate, sulfur, and carbon dioxide. The vast majority of aromatics production from coal occurs in Eastern Europe, India, and Japan.[75]

The combination of chemical and fuel production results in a type of synergism that can benefit the economics of both end products. Thus, some chemicals that cannot be produced economically as primary products from coal may be obtained as revenue-generating coproducts that enhance the overall economics of fuel production. This concept has prompted efforts at Great Plains and Sasol to increase the recovery of chemical by-products from these facilities. At Great Plains, ammonia and sulfur now are recovered, but facilities are being added to recover phenol from gasification liquids and a krypton/xenon mixture from the air separation plant. R&D efforts are in progress to enable economical recovery of other by-products.

In recent years several commercial plants have been constructed for conversion of coal to synthesis gas for chemical manufacturing. These include the Eastman acetic anhydride plant, the Ube ammonia plant, and the SAR oxo chemicals plant. Of these, the Eastman plant is the only one still operating exclusively on coal. Feedstock changes at the other plants illustrate the vulnerability of coal conversion processes to a changing economic climate. The fact that the Eastman process remains competitive under changing conditions is due to a set of special circumstances that favor a coal-based process. The success of the Eastman Chemicals from Coal complex demonstrates that synthesis gas from coal is a viable feedstock for some industrial chemicals, given a specific set of conditions.

Clearly, the slate of chemicals produced from coal-derived synthesis gas will expand as new technologies are developed, and supplies of petroleum and natural gas dwindle. The most likely such chemicals are those for which existing processes have been demonstrated but which presently lack economic merit. Relatively small improvements in technology, shifts in feedstock availability, capital costs, or political factors could enhance the viability of coal-based processes for the production of methanol, ethanol, and higher alcohols, vinyl acetate, ethylene glycol, carboxylic acids, and light olefins.

Developing technologies presently exist for the partial oxidation of coal to organic acids, for plasma pyrolysis of coal to produce acetylene and other small molecules,[76] and for the manufacture of additional chemicals such as acetaldehyde from synthesis gas.[77] Also, the use of coal tar for the manufacture of high-temperature engineering plastics[78] and bioconversion of coal to aromatics, organic acids, alcohols, and methane[79] have been reported. Innovative strategies are being considered that offer potential for reducing the costs of energy and chemicals produced

from coal by taking avantage of waste disposal credits through gasification of coal mixed with carbon-containing waste streams.[80]

Mild gasification is another emerging technology being pursued. The objective of this technology is to produce an optimum combination of products, including fuel and chemicals, at conditions considerably less severe than required for synthesis gas generation.[81] The optimum product mix varies with the coal type. Benzene and phenol are large-volume chemicals potentially available from mild gasification of eastern coals, and several high-value carbon products are available from the char produced using western coals.[82]

Indirect Coal Liquefaction

In indirect coal liquefaction, a synthesis gas first is produced in a suitable gasifier by reaction with oxygen and steam. Then this gas is treated to remove sulfur compounds, and the ratio of CO to H_2 is fixed by means of the water-gas shift reaction. The purified, adjusted synthesis gas then is reacted to produce the desired liquid products.

Fischer Tropsch—The SASOL Complex. Fischer-Tropsch synthesis is the best-known indirect liquefaction method. It was developed in 1925 by German chemists Franz Fischer and Hans Tropsch. In this process, synthesis gas is reacted over an iron or cobalt catalyst at 350 psi and 330°C to produce largely aliphatic hydrocarbons, including gasoline, jet fuel, diesel oil, middle distillates, heavy oil, and waxes. The Germans used this technology during World War II to produce nearly 15,000 barrels/day of military fuels.

The South Africans have applied the Fischer-Tropsch technology in their SASOL facility. The Sasol (South African Coal, Oil and Gas Corporation) installation is the world's largest synfuel plant based on coal. It converts an inferior high-ash (35%) coal into a broad spectrum of products, which may include ethylene, propylene, butylene, gasoline, and long straight-chain alcohols and hydrocarbons. The Sasol plant has been operating for over 35 years.

South Africa has substantial deposits of coal but no oil and obtains 75 percent of its energy from coal. The plant at Sasolburg is located on a huge coal field. SASOL I, the existing synthetics plant, is part of a versatile chemical complex that supplies much of South Africa's needs for materials such as nitrogen fertilizers, plastics, and pipeline gas. Consumption of coal at Sasol is 5 million tons/year. Coal is gasified in Lurgi reactors with steam and oxygen under pressures of about 20 atmospheres. Principal products are H_2, CO, CO_2, and CH_4. The key components are H_2 and CO, and their ratio can be adjusted by changing the amounts of H_2O and O_2 used in the gasification. Constituents present in addition to the major gaseous products include NH_3, H_2S, and other sulfur-containing substances. The raw products are purified by cooling the gas, condensing out water and phenolic-type substances, and passing the remaining gas through three absorption trains using very cold methanol. This removes essentially everything except H_2, CO, and CH_4. Sulfur in the purified gas amounts to less than 1×10^{-8}. This guarantees relative freedom from sulfide poisoning of catalysts and a low sulfur content in the final products. A flow sheet is shown in Fig. 17.36.

In October 1973, a large-scale plant, SASOL II, designed primarily to produce gasoline, was authorized by South Africa. In 1978, another major sythetic fuel plant was started. These two new plants were designed for zero emissions of liquid wastes and were intended to produce over 100,000 barrels of coal-derived liquids/day. The total cost of these projects was over $7 billion.

In 1980, Sasol introduced a second generation of its Fischer-Tropsch technology (the Synthol process), and in 1989 a third generation. Each new generation has utilized improved engineering design and has enhanced the utilization of the chemicals generated. At Sasol, the process is operated to produce primarily gasoline and diesel fuels, but it can be modified by using different catalysts and reaction conditions to ehance the yield of oxygenated products that can be separated for chemical uses.[74,83] Some of the oxygenates available

Sasol's Two Main Coal-Into-Oil Processes

Coal gasification and purification flow diagram

(a)

Fischer-Tropsch Synthesis Flow Diagram

(b)

Fig. 17.36. Sasol's two main coal-into-oil processes. (*Reprinted from* Encyclopedia of Chemical Processing Design, *pp. 310, 320, by courtesy of Marcel Dekker, Inc., 1979.*)

from the process are methanol, ethanol, propanol, butanone, and higher alcohols and ketones. Fischer-Tropsch olefins can be separated and used for production of oxo chemicals, synthetic rubber, and alkylbenzene sulfonate detergents. Hydrocarbon waxes produced in the reactors have found a variety of uses. Also, by-products from the Sasol Lurgi coal gasifiers are recovered for chemical and solvent applications. These products include phenol, cresols, toluene, xylenes, ammonia, and sulfur. A recent addition to the spectrum of chemical products from Sasol is polypropylene. Also, Fischer-Tropsch ethane now is being cracked to supplement ethylene production for sale to polyethylene producers. Additional work is in progress to evaluate the recovery of organic acids from aqueous waste streams.

Methanol. Methanol is a water-soluble, low molecular weight alcohol that may be of increasing importance as a low-sulfur fuel, a chemical feedstock, and perhaps an intermediate for the production of gasoline. The synthesis of methanol is accomplished by the catalytic conversion of synthesis gas containing two moles of hydrogen for each mole of carbon monoxide. Methanol synthesis is widely practiced in industry on a commercial scale and will not be discussed further here.

Chemicals from Coal at Eastman Chemical Company

In 1983 Eastman Chemical Company (Division of Eastman Kodak Company) became the first U.S. manufacturer to produce a modern generation of industrial chemicals, with acetic anhydride as the key product, using coal as the feedstock.[84] This production was made possible by successful completion of facilities at Tennessee Eastman Division. in Kingsport, Tennessee. The start-up of this chemical plant was the result of more than a decade of work done to identify, develop, and assemble the numerous technologies needed for a viable commercial venture.

Commercial facilities constructed include a coal gasification plant for synthesis gas manufacture, raw gas cleanup and separation facilities, a sulfur recovery unit, a coal-fired steam plant, and chemical plants to produce methanol, methyl acetate, acetic acid, and acetic anhydride. In the 1980s, these plants gasified approximately 900 tons of coal/day to produce annually over 500 million lb of acetic anhydride, 150 million lb of acetic acid, 390 million lb of methyl acetate, 365 million lb of methanol, and 15 million lb of sulfur. An expansion in 1991 more than doubled the capacity for methyl acetate, acetic acid, and acetic anhydride. Approximately one billion lb of acetic anhydride, one of Eastman's key raw materials, is used each year in making photographic film base, Tenite cellulosic plastics, Estron acetate for filter tow, textile yarns, and coatings chemicals.

A flow sheet of Eastman's chemicals from coal plant is shown in Fig. 17.37. A simplified description of the main plants in the complex follows.

A. *Coal gasification.* In the mid-1970s Eastman engineers and scientists studied modern coal gasification technologies and the modifications necessary to produce chemical feedstocks. After all factors were considered, the Texaco coal gasification process was chosen to provide synthesis gas from the local coal. In this process, oxygen and coal slurry are introduced into the gasifier at elevated temperature and pressure. The Texaco-licensed process operates at sufficiently high temperatures to avoid producing environmentally undesirable by-products. The slag from this process is suitable for disposal in nonhazardous landfills. The energy recovery arrangement produces medium- and low-pressure steam, which is well suited for use by other processes in the complex. The product gas is mainly carbon monoxide and hydrogen, making this unit particularly suitable for the production of chemical feedstocks. Because the raw gas is generated at elevated pressures, the size of the gas cleanup equipment is reduced, and the need to compress the product gas prior to its use as a chemical feedstock is minimized.

Fig. 17.37. Overall block flow diagram for coal gasification–acetic anhydride complex.

The gasifier product is scrubbed with water to cool the gas and remove any ash particles. A portion of the gas stream is sent to a water–gas shift reaction to increase the hydrogen content. Finally, low-pressure steam is generated as the product gas is cooled before additional purification. The full recovery of this heat increases the thermal efficiency of the process and is important for economical operation of the chemical processes that utilize the steam.

B. *Synthesis gas refinement.* Following the gas scrubbing, shift reaction, and cooling operations associated with the gasifier, the product gas streams are treated to remove hydrogen sulfide in a Rectisol process by utilizing Linde, A. G., technology. In this process, the sulfur in the coal, which is converted to hydrogen sulfide during the gasification process, and carbon dioxide are absorbed by a cold methanol wash. Further processing cryogenically purifies the carbon monoxide and hydrogen streams in a "cold box," also provided by Linde. The cryogenic separation provides a carbon monoxide feed for the acetic anhydride plant and a hydrogen-rich stream suitable for methanol production. These processes are uniquely integrated to improve the overall thermal efficiency of their operation.

C. *Sulfur recovery.* The hydrogen sulfide removed from the product gas streams is further processed into elemental sulfur. Conversion of the sulfide to elemental sulfur results in 99.7 percent of the sulfur originally contained in the coal being recovered and sold as a by-product. This conversion is accomplished using the Claus process followed by a Shell Claus Off-Gas Treating unit (SCOT).

D. *Methanol production.* Methanol is produced from carbon monoxide, carbon dioxide, and hydrogen. The proper gas-feed composition for methanol production is formed by combining the hydrogen-enriched syngas from the shift reactor and the hydrogen–carbon monoxide stream from the gas separation unit. The refined syngas streams, which contain a proper combination of hydrogen and carbon oxides, are fed to a low-pressure, gas-phase catalytic methanol process that is very energy-efficient, both in the method of control of the reactor and in subsequent refinement of the methanol. This process and the proprietary items of equipment were provided by Lurgi, A. G.

E. *Gasification system venting.* Additional efficiency was achieved by providing for combustion of process vent gas streams having fuel value in a nearby powerhouse. Emergency venting of the syngas is directed to one of the two flare systems provided to

avoid the development of undesirable or hazardous conditions during process upsets.

F. *Methyl acetate production.* Methanol made from synthesis gas is reacted with by-product acetic acid from cellulose esters manufacturing to form methyl acetate in an Eastman patented process (U.S. Patent 4,435,595). The principal component of this process is a reactor-distillation column in which acetic acid and methanol flow countercurrently, reacting and flashing at each stage. The acetic acid serves the dual functions of reactant and extractive agent to remove water and methanol from their methyl acetate azeotropes.

G. *Production of acetic anhydride and acetic acid.* In the final step of the process, purified carbon monoxide from the gas separation plant is reacted with methyl acetate to form acetic anhydride by using a proprietary catalyst system and process developed by Eastman (U.S. Patent 4,374,070). Part of the acetic anhydride is reacted with methanol to coproduce acetic acid. The methyl acetate resulting from this reaction is recarbonylated in the acetic anhydride reactor system. The acetic acid and acetic anhydride are refined in the distillation section of this plant. A reactor liquid purge stream removes high molecular weight by-products, and a catalyst recycle process recovers and regenerates the catalyst from the purge stream (U.S. Patent 4,388,217).

The Eastman Chemicals from Coal Plant has been demonstrated to provide an economically attractive process for commercial production of acetic anhydride. It is instructive to recapitulate some of the factors that contribute to this situation.

A. *Plant configuration.* (1) The hydrogen-to-carbon monoxide ratio of the syngas is well suited to combined production of methanol and acetic anhydride. Although the hydrogen is not needed for production of acetic anhydride via carbonylation, it is used to decrease the hydrogen deficiency in the feed gas and thus minimizes use of the water–gas shift reactor. (2) The methyl acetate and acetic anhydride plants are highly energy-efficient, and they use medium and low-pressure steam produced by cooling the gasifier product. (3) The acetic anhydride is used in existing processes to make cellulose acetate. The reaction produces by-product acetic acid that can be recycled to the methyl acetate plant to react with methanol. This combination of processes obviated the need to construct an acetic acid plant.

B. *Coal feedstock advantages.* (1) The operating cost is less than for a natural gas or petroleum-based plant. (2) The coal cost is closely related to mining and delivery costs and is less subject to unstable political environment than that of petroleum. (3) The plant location near Appalachian coal fields minimizes delivery costs. (4) Coal gasifiers can use less expensive high-sulfur coal, and the sulfur can be recovered and sold to further improve the economics.

C. *Environmental advantages.* (1) No tar or other hydrocarbon waste streams are produced in the high-temperature (over 2500°F) gasifier. (2) The molten slag is converted to inert granular pellets that qualify for nonhazardous landfill. (3) Wet grinding of the coal to form a slurry eliminates dust problems. (4) Because the sulfur in the coal is converted to hydrogen sulfide in the gasifier, more than 99.7 percent is eventually recovered as elemental sulfur. (5) Aqueous streams containing hydrocarbons from the chemical plants are recycled to the gasifier for slurry and quench water makeup. (6) Tail-gas streams from the chemical plants are burned for heat recovery in existing boilers and furnaces.

Methanol to Gasoline—The Mobil Process. Mobil Research and Development Corporation, under a contract with the U.S. Department of Energy, developed a process that catalytically dehydrates and polymerizes methanol to produce a high octane unleaded gasoline. The catalyst is one of a new family of synthetic zeolites designated ZSM-5 by Mobil. These new zeolites have a unique channel structure, different from previously known

Fig. 17.38. Mobil's fluid-bed MTG process uses a unique zeolite catalyst in converting methanol directly to high octane, unleaded gasoline. (*Courtesy The Pace Company, Denver, CO.*)

wide-pore (9–10 Å in diameter) and other narrow-pore (5 Å) zeolites. One of these new zeolites converts methanol into a mixture of hydrocarbons corresponding to high-quality gasoline. The mass balance shows 0.44 lb of gasoline and 0.56 lb of water from each pound of methanol, and the heating value of the gasoline is 95 percent of that in the methanol feed. A schematic diagram of the process is shown in Fig. 17.38. Crude methanol is vaporized at the bottom of the reactor and passes through the dense fluidized bed at 775°F and 25 psig. The methanol is converted to hydrocarbons and water. The catalyst is removed from reaction products in a disengager section at the top of the reactor, the reactor effluent is condensed, and the water and hydrocarbon products are separated. To make additional gasoline, propene and butenes can be alkylated with isobutane by conventional petroleum technology. The reactor is 25 feet long and 4 inches in diameter.

Portions of the catalyst powder are periodically removed from the reactor and regenerated with air. The regenerated catalyst is transferred back to the reactor. Small amounts of carbon monoxide, carbon dioxide, and coke are formed as by-products.

REFERENCES

1. *Annual Energy Review*, Energy Information Administration, U.S. Department of Energy, DOE/EIA-0384 (90), 1990.
2. Given, P. H., Chapter 12 in *Coal, Its Formation and Composition*, W. Francis (Ed.), Edward Arnold, Ltd., London, 1961.
3. Wender, I., Heredy, L. A., Neuworth, M. B., and Dryden, I. G. C., 1981. "Chemical Reactions and the Constitution of Coal," in *Chemistry of Coal Utilization*, M. A. Elliott (Ed.), Second Supplementary Volume, pp. 425–522, John Wiley & Sons, New York, 1981.
4. Gavalas, G. R., "Coal Pyrolysis," in *Coal Science and Technology* Vol. 4, Elsevier, 1982.
5. Horwitch, M., Schuller, F., Stobaugh, R., and Yergin, D. (Eds.), "Coal: Constrained Abundance," in *Energy Future: Report of the Energy Project at the Harvard Business School*, pp. 79–107, Random House, New York, 1979.
6. Khan, M. R., *Fuel*, **68** (11), 1439–1449 (1989).
7. Leonard, J. W. (Ed.), *Coal Preparation*, 5th ed., American Institute of Mining, Metallurgical and Petroleum Engineers, Inc., New York, 1992.
8. Osborne, D. G., *Coal Preparation Technology*, Graham & Trotman, 1989.
9. Killmeyer, R. P., Hucko, R. E., and Jacobsen, P. S., "Interlaboratory Comparison of Advanced Froth Flotation Processes," preprint No. 89–137, Society of Mining Engineers, Littleton, CO (1989).
10. Parekh, B. K., Groppo, J. G., Stotts, W. F., and Bland, A. E., "Recovery of Fine Coal from Preparation Plant Refuse Using Column Flotation," *Column Flotation '88*, K.V.S. Society (Eds.), Society of Mining Engineers, Inc., Littleton, CO, 1988.
11. Parekh, B. K., and Bland, A. E., "Fine Coal and Refuse Dewatering—Present State and Future Consideration," *Flocculation and Dewatering*, B. M. Moudgil and B. J. Scheiner (Eds.), Engineering Foundation, New York, 1989.
12. Khan, M. R., and Kurata, T., in "The Feasibility of Mild Gasification of Coal: Research Needs," DOE/METC-85/4019, NTRS/DE85013625, 73 pp. (1985).
13. Probstein, R. F., and Hicks, R. E., *Synthetic Fuels*, McGraw-Hill, New York, 490 pp., 1982.
14. Seglin, L., and Bresler, S. A., "Low-Temperature Pyrolysis Technology," in *Chemistry of Coal Utilization*, M. Elliott (Ed.), Second Supplementary Volume, pp. 785–846, John Wiley & Sons, New York, 1981.
15. Howard, J. B., "Fundamentals of Coal Pyrolysis and Hydropyrolysis," in *Chemistry of Coal Utilization*, M. A. Elliott (Ed.), Second Supplementary Volume, pp. 665–784, John Wiley & Sons, New York, 1981.
16. Solomon, P. R., and Serio, M. A., "Evaluation of Coal Pyrolysis Kinetics," in *Fundamentals of Physical Chemistry of Pulverized Coal Combustion*, J. Lahaye and G. Prado (Eds.), Martinus Nijhoff Publishers, 1987.
17. Khan, M. R., "Thermoplastic Properties of Coal Pyrolysis at Elevated Pressures: Effects of Experimental Variables, Inorganic Additives, and Preoxidation," Ph.D. thesis, The Pennsylvania State University, 333 pp. (1985)
18. Anthony, D. B., and Howard, J. B., "Coal Devolatilization and Hydrogasification," *AIChE Journal* **24** (2), 625–656 (1976).
19. Solomon, P. R., Serio, M. A., and Suuberg, E. M., "Review of Coal Pyrolysis: Experiments, Kinetics Rates and Mechanisms," *Progress in Energy and Combustion Science*, in press (1992).
20. Gryaznov, N. S., and Petrov, V. K., *Coke and Chemistry*, pp. 27–32, USSR, (Coal Tar RA Translation), 1960.
21. Given, P. H., and Yarzab, R. F., Chapter 20, "Analysis of the Organic Substance of Coals: Problems Posed by the Presence of Mineral Matter," in *Analytical Methods for Coal and Coal Products*, Vol. II, C. Karr, Jr. (Ed.), pp. 3–41, Academic Press, Inc., New York, 1978.
22. Walker, P. L., Jr., Davis, A., Verma, S. K., Rivera-Utrilla, J., and Khan, M. R., "Interaction of Gases, Vapors, and Liquids with Coals and Minerals at Ambient Temperatures—Their Use to Characterize Porosity," The Pennsylvania State University, DOE-30013-19, Under Contract No. DE-AC22-80PC30013. (1984).
23. Khan, M. R., *Fuel*, **66**, 1626 (1987).
24. Walker, P. L., Jr., "Coal Gasification: University Contribution of the Story of Devolatilization, Catalysis, and Active Sites," in *Symposium on Advances in Coal Utilization Technology*, pp. 237–271, IGT, CONF-790598, 1979.
25. Shinner, R., Fourth Technical Meeting of COGARN, presented in "Coal Gasification: Direct Applications and Synthesis of Chemicals and Fuels; A Research Needs Assessment," by Penner, S. S., and others for US DOE, Office of Energy Research, DOE Contract No. DE-AC01-85ER30076, NTIS-PR-360 (1987).
26. Penner, S. S., et al., "Coal Gasification: Direct Applications and Synthesis of Chemicals and Fuels; A Research Needs Assessment," DOE Contract No. DE-AC01-85ER30076, NTIS-PR-360 (1987).
27. Spencer, D. F., Alpert, S. B., and Gluckman, M. J., "Integrated Coal Gasification Combined Cycles (IGCC) an Emerging Commercial Option for the Power Industry," paper presented at 1985 ACS Meeting, CA, 1985.
28. Spencer, D. F., 1986, "The Commercial Implications of the Cool Water Project for the Electric Power Industry," EPRI, Palo Alto, CA (1986).
29. Spencer, D. F., Alpert, S. B., and Gilman, H. H., *Science*, **232**, 609 (1986).
30. Smoot, L. D., "Coal and Char Combustion," in *Fossil Fuel Combustion*, W. Bartok and A. F. Sarofim (Eds.), John Wiley & Sons, New York, 1991.

31. Corcoran, E., "Cleaning Up Coal," *Scientific American*, **264** (5), 106 (1991).
32. Haggin, J., "Coal," *Chemical and Engineering News*, p. 32, (June 17, 1991).
33. Schobert, H. H., *Coal: The Energy Source of the Past and the Future*, American Chemical Society, Washington, DC, 1987.
34. Seikel, G, R., "Coal-Fired Open-Cycle MHD Plants," in *The Science and Technology of Coal Utilization*, B. R. Cooper, and W. A. Ellingson (Eds.), Plenum Press, New York, 1984.
35. Ceely, F. J., and Daman, E. L., "Combustion Process Technology," in *Chemistry of Coal Utilization*, M. A. Elliott (Ed.), Wiley-Interscience, New York, 1981.
36. La Nauze, R. D., "Combustion in Fluidized Beds," in *Advanced Combustion Methods*, F. J. Weinberg (Ed.), Academic Press, New York, 1986.
37. Smoot, L. D., and Smith, P. J., *Coal Combustion and Gasification*, Plenum, New York, 1985.
38. Miller, B. G., "Coal–Water Slurry Fuel Utilization in Utility and Industrial Boilers," *Chemical Engineering Progress*, **85** (3), 29 (1989).
39. Bartok, W., Lyon, R. K., McIntyre, A. D., Ruth, L. A., and Sommerlad, R. E., "Combustors: Applications and Design Considerations," *Chemical Engineering Progress*, **84** (3), 54 (1988).
40. Singer, J. G. (Ed.) *Combustion-Fossil Power Systems*, Combustion Engineering, Inc., Windsor, CT, 1991.
41. Podolski, W. F., "Fluidized-Bed Combustion," in *The Science and Technology of Coal Utilization*, B. R. Cooper, and W. A. Ellingson, (Eds.), Plenum Press, New York, 1984.
42. Dry, R. J., and La Nauze, R. D., "Combustion in Fluidized Beds," *Chemical Engineering Progress*, **86** (7), 31 (1990).
43. U.S. Department of Energy, "Clean Coal Technology Demonstration Program," DOE/FE-0219P (Feb. 1991).
44. Wu, W. R. K., and Storch, H. H., "Hydrogenation of Coal Tar," U.S. Bureau of Mines Bulletin No. 633 (1968).
45. Donath, E. E., and Hoering, M., "Early Coal Hydrogenation Catalysis," *Fuel Proc. Tech.*, **1**, 3–20 (1977).
46. Ministry of Fuel and Power, *Report on the Petroleum and Synthetic Oil Industry of Germany*, Ministry of Fuel and Power, His Majesty's Stationery Office, London, 1947.
47. Neavel, R. C., "Liquefaction of Coal in Hydrogen-Donor and Non-donor Vehicles," *Fuel*, **55**, 237–242 (1976).
48. Derbyshire, F., and Stansberry, P., "Comments on the Reactivity of Low-Rank Coals in Liquefaction," *Fuel*, **66**, 1741–1742 (1987).
49. Holmes, D. F., Jr., Mitchell, J., and Farthing, S., "A Final Report of Kentucky's Role in the H-Coal Pilot Plant Project," Kentucky Energy Cabinet (Aug. 1983).
50. DOE, "Coal Liquefaction: A Research and Development Needs Assessment," DOE Coal Liquefaction Research Needs (COLIRN) Panel Assessment, Final Report, Vol. 2, DOE/ER-0400 (Mar. 1989).
51. Derbyshire, F., "Coal Liquefaction," in *Ullmann's Encyclopedia of Industrial Chemistry*, Vol. A7, VCH Publishers, New York, 1986.
52. Furlong, L. E., Effron, E., Vernon, L. W., Wilson, E. L., "Coal Liquefaction—the Exxon Donor Solvent Process," presented at the AIChE National Meeting, Los Angeles, CA, Nov. 18, 1975.
53. Maa, P. S., Trachte, K. L., and Williams, R. D., "Solvent Effects in Exxon Donor Solvent Coal Liquefaction," presented at the ACS National Meeting, New York, Aug. 23–28.
54. Neavel, R. C., "Exxon Donor Solvent Liquefaction Process," *Phil. Trans. R. Soc. Lond. A*, **300**, 141–156 (1981).
55. Tomlinson, G. C., Gray, D., Neuworth, M. B., Talib, A., "The Impact of Rank-Related Coal Properties on the Response of Coals to Continuous Direct Liquefaction Processes," Sandia Contractors Report, SAND85-7238 (Oct. 1985).
56. Derbyshire, F., "Catalysis in Coal Liquefaction: New Directions for Research," IEACR/08, IEA Coal Research, London (1968).
57. EPRI, "Two-Stage Coal Liquefaction Integrated Configurations: The Advanced Coal Liquefaction R&D Facility, Wilsonville, Alabama," EPRI GS-7293 (May 1991).
58. Neuworth, M. B., and Moroni, E. C., "Development of an Integrated Two Stage Coal Liquefaction Process," *Fuel Proc. Tech.*, **8**, 231–239 (1984).
59. Gollakota, S. V., Lee, J. M., and Davies, O. L., "Process Optimization of Close-Coupled Integrated Two-stage Liquefaction by the Use of Cleaned Coals," *Fuel Proc. Tech.*, **22**, 205–216 (1989).
60. El Sawy, A., Gray, D., Talib, A., and Tomlinson, G., "A Techno-Economic Assessment of Recent Advances in Direct Coal Liquefaction," Sandia Contractor Report SAND86-7103 (June 1986).
61. Kulik, C., Lebowitz, H., Weber, W., and Johnson, T. W., "Coal Liquefaction with Selective Heavy Recycle," presented at the ACS National Meeting, New York, Aug. 23–28, 1981.
62. Nalitham, R. V., Lee, J. M., Lamb, C. W., and Johnson, T. W., "Two-Stage Coal Liquefaction Process Performance with Close-Coupled Reactors," *Fuel Proc. Tech.*, **17**, 13–27 (1987).
63. Rosenthal, J. W., Dahlberg, A. J., Kuehler, C. W., Cash, D. R., and Freedman, W., "The Chevron Coal Liquefaction Process (CCLP)," *Fuel*, **61**, 1045–1049 (1982).
64. Lee, J. M., Davies, O. L., Pinkston, T. E., Gough, J. R., "Recent Developments in Coal Liquefaction at Wilsonville," *Proceedings: 13th Annual EPRI Conference on Fuel Science and Conversion*, EPRI GS-6219, Jan. 1989.
65. Direct Liquefaction Contractor's Review Meeting, *Proceedings*, Pittsburgh, PA, 24–26, 1990.

66. Gray, M. D., Kimber, G. M., Moore, S. A., "British Coal's Liquid Solvent Extraction (LSE) Process," *Proceedings: 14th Annual EPRI Conference on Fuel Science*, EPRI GS-6827, May 1990.
67. Kimber, G., "Petrol and Diesel from Coal: British Coal's Liquefaction Process," *Energeia*, **2**, 1–2 (1991).
68. Derbyshire, F. J., "Role of Catalysis in Coal Liquefaction Research and Development," *Energy and Fuels*, **3**, 273–277 (1989).
69. Wender, I., in *Chemicals from Coal: New Processes*, K. R. Payne (Ed.), *Critical Reports on Applied Chemistry*, Vol. 14, Chapter 3, John Wiley & Sons, Chichester, 1987.
70. Cornils, B., in *Chemicals from Coal: New Processes*, K. R. Payne (Ed.), *Critical Reports on Applied Chemistry*, Vol. 14, Chapter 5, John Wiley & Sons, Chichester, 1987.
71. Pruett, R. L., *Science*, **211**, 11–16 (1981).
72. Shreve, R. N., *Chemical Process Industries*, Chapter 5, McGraw-Hill, New York, 1967.
73. Spitz, P. H., *CHEMTECH*, **19**, 92–100 (1989).
74. Schultz, H., *Pure and Appl. Chem.*, **51**, 2225–2241 (1979).
75. Janardanarao, M., *Journal of Scientific and Industrial Research*, **43**, 649–659 (1984).
76. Anderson, L. L., and Tillman, D. A., *Synthetic Fuels from Coal*, Wiley Interscience, New York, 1979.
77. Moloy, K. G., and Wegman, R. W., *Organometallics*, **8**, 2883 (1989).
78. Jedlinski, Z. J., and Gaik, U., *Polimery*, **29**, 424 (1984).
79. Haggin, J., *Chemical and Engineering News*, **67** (7), 28 (1989).
80. *Coal and Synfuels Technology*, **11** (16), 1 (1990).
81. Haggin, J., *Chemical and Engineering News*, **68** (21), 33 (1990).
82. Sondreal, E. A., Ness, R. O., Timpe, R. C., Knudson, C. L., Jha, M. C., Berggren, M. H., and Sinor, J. E., *Proc. Sixth Annual Int. Pittsburgh Coal Conf.*, Vol. 2, 785–794, Sept. 25–29, 1989.
83. Dry, M. E., in *Chemicals From Coal: New Processes*, K. R. Payne (Ed.), *Critical Reports on Applied Chemistry*, Vol. 14, Chapter 4, John Wiley & Sons, Chichester, 1987.
84. Agreda, V H., *Chemtech*, **18**, 250–253 (1988).
85. Agreda, V. H., *Fuel*, **69**, 132 (1990).
86. Arthur, J. R., and Wadsworth, K. D., "The Effect of Inorganic Compounds on the Behavior of Coals and Cokes," *BCURA Progress in Coal Science*, pp. 396–410 (1950).
87. Crynes, B. L., 1981. "Processing Coal Liquefaction Products," in *Chemistry of Coal Utilization*, M. A. Elliott (Ed.), Second Supplementary Volume, pp. 1991–2070, John Wiley & Sons, New York, 1981.
88. Dryden, I. G. C., "Chemical Constitution and Reactions of Coal," in *Chemistry of Coal Utilization, Supplementary Volume*, H. H. Lowry, (Ed.), pp. 223–295, John Wiley & Sons, New York, 1963.
89. Huxtable, W. H., *Coal Tar Fuels: Their Production, Properties, and Application*, Association of Tar Distillers, London, Jan. 1961.
90. Oara, J. B., et al., "Petrochemical Feedstocks from Coal," *Chem. Eng. Prog.*, pp. 64–72, June 1977.
91. Khan, M. R., *Fuel*, **68** (12), 1522–1531 (1989).
92. Lee, J. M., and Cantrell, C. E., "Effect of Coal Rank and Quality on Two-Stage Liquefaction," *Proceedings: 14th Annual EPRI Conference on Fuel Science*, EPRI GS-6827, May 1990.
93. Lowry, H.H. (Ed.), *Chemistry of Coal Utilization*, Supplementary Volume, pp. 232–295, John Wiley & Sons, New York (1962).
94. Synthetic Fuels Assn., Inc., "Coal Gasification Systems: A Guide to Status, Applications and Economics," EPRI AP-3109, Project 2207, (June 1983).
95. O'Hara, J. B. et al., "Project POGO—A Coal Refinery," *Chem. Eng. Prog.*, pp. 46–63, August 1978.
96. Weber, W., and Stewart, N., "Direct Coal Liquefaction," *EPRI Journal* **12** (1), 40–41 (1987).
97. Wilson, P. J., Jr., and Clendenin, J. D., "Low-Temperature Carbonization," in *Chemistry of Coal Utilization*, H. H. Lowry (Ed.), Supplementary Volume, pp. 395–460, John Wiley & Sons, New York, 1963.
98. *Conference on Chemical Feedstocks Alternatives*, American Institute of Chemical Engineers and National Science Foundation, Houston, Texas, October 2–5, 1977.
99. Funk, J. E., "Coal Technology," in *Riegel's Handbook of Industrial Chemistry*, 8th ed., J. A. Kent (Ed.), Van Nostrand Reinhold, New York, 1983.

PATENTS

"Reactive Distillation Process for the Production of Methyl Acetate," V. H. Agreda and L. R. Partin, U.S. Patent No. 4,435,595, Mar. 6, 1984.

"Preparation of Acetic Anhydride," T. H. Larkins, S. W. Polichnowski, G. C. Tustin, and D. A. Young, U.S. Patent 4,374,070, Feb. 15, 1983.

"Process for the Recovery of Catalyst Values," R. T. Hembre and S. L. Cook, U.S. Patent 4,388,217, June 14, 1983.

ADDITIONAL REFERENCES

Cassidy, E. M., *Elements of Practical Coal Mining*, Port City Press, Inc., Baltimore, MD, 1973.

Coal Production 1989, Energy Information Administration, U.S. Department of Energy, Washington, DC 20585.

Hustruled, W. A. (Ed.), *Underground Mining Methods Handbook*, Society of Mining Engineers, The American Institute of Mining, Metallurgical and Petroleum Engineers, 1982.

18

Rubber

T. C. Bouton*

INTRODUCTION

The word "rubber" immediately brings to mind materials that are highly flexible and will snap back to their original shape after being stretched. In this chapter a variety of materials are discussed that possess the characteristics describing this class of materials. Originally rubber meant the gum collected from a tree growing in Brazil. The term "rubber" was coined for this material by the English chemist Joseph Priestley, who noted that it was effective for removing pencil marks from paper. Today, in addition to Priestley's natural product, many synthetic materials are made that possess these characteristics and many other properties. The common features of these materials are that they are made up of long polymer chains that are amorphous (not crystalline), they have a high degree of flexibility such that they do not stiffen until temperatures well below ambient are reached, and the chains can form a three-dimensional network in some manner.

It is difficult, indeed almost impossible, to visualize a modern society without rubber products. The development of synthetic rubber began out of the need for countries to establish independence from natural products that grew only in tropical climates. In times of conflict the natural product might not be available, and its loss would seriously threaten national security. Synthetic rubber, then, became a strategic concern during World Wars I and II.[1] Beyond the security issue, the need for materials with better performance also provided a strong impetus for the development of new rubbery materials. In particular, improvements in oil resistance, high temperature stability, and oxidation and ozone resistance were needed.[2] Research today is driven to develop materials with even better performance in these areas.

The demand for rubber was estimated to be 15.845 million metric tons for 1990.[3] Of this demand, 67 percent is synthetic rubber of various kinds. Rubber has been classified by use into general purpose and specialty. The major general purpose rubbers are natural rubber, styrene-butadiene rubber, butadiene rubber, isoprene rubber, and ethylene-

*Firestone Synthetic Rubber & Latex Co., Division of Bridgestone/Firestone, Inc., Akron, OH.

propylene rubber. These rubbers are used in tires, mechanical goods, and similar applications. Specialty elastomers provide unique properties such as oil resistance or extreme heat stability. Although this differentiation is rather arbitrary, it tends also to classify the polymers according to volumes used. Styrene-butadiene rubber, butadiene rubber, and ethylene-propylene rubber account for 75 percent of all synthetic rubber consumed.

The 1990 estimated synthetic rubber capacity by type is presented in Table 18.1, and Table 18.2 lists the U.S. rubber consumption by type in 1989.

TABLE 18.1 Synthetic Rubber Production by Type (Excludes Centrally Planned Economy Countries)

Type	*Description*	*Capacity Metric Tons*
SBR-solid	Styrene/butadiene rubber (emulsion plus solution)	2,430,000
SBR-latex	Styrene/butadiene emulsion rubber–latex	315,000
BR	Butadiene rubber	968,000
IR	Isoprene rubber	129,000
CR	Chloroprene (or neoprene) rubber	308,000
EPDM	Ethylene/propylene terpolymer	573,000
NBR	Nitrile or acrylonitrile/ butadiene rubber	262,000
Others		595,000

Worldwide Rubber Statistics 1990, International Institute of Synthetic Rubber Products, Inc.

TABLE 18.2 U.S. Rubber Consumption, 1989 (Metric Tons)

Total new rubber	2,917,836
Natural	866,870
Total synthetic	2,050,966
SBR	769,579
Nitrile	65,645
Polybutadiene	374,793
EPDM	288,855
Butyl polyisoprene, polychloroprene, silicone, and others	612,094
% Natural	29.7
% Synthetic	70.3

"RMA Statistical Report," Rubber Manufacturers Association, 1400 K Street, NW, Washington, DC 20005.

RUBBER CONCEPTS

Several key principles (outlined below) greatly help interested persons in understanding the performance of elastomeric materials. This outline should present these concepts well enough for use in the following discussions of specific polymer types. However, the reader should consult the references for a more complete understanding of these principles. The concepts are classified as those that relate to polymer structure, those that relate directly to physical properties, and those that relate to the use of the material, as listed below.

1. Polymer structure
   a. Macrostructure
      - Molecular weight
      - Molecular weight distribution
      - Branching
   b. Microstructure
   c. Copolymer composition
   d. Network structure
2. Rubber properties
   a. Glass transition temperature
   b. Crystallinity
3. Rubber use
   a. Compounding
   b. Processing

POLYMER STRUCTURE

Macrostructure

Molecular Weight. The single most important property of any polymer is the size or length of the molecule. The polymerization process consists of combining many of the simple monomer molecules into the polymer chain. Most of the monomers used to produce rubbers are either gases or low-viscosity liquids under normal conditions; upon polymerization they form liquids whose viscosity increases to extremely high values as the chain length is increased. At very low chain length, this increase is linear with molecular weight until the chains are long enough to become entangled. Above the entanglement molecular weight, the viscosity increases to the 3.4–3.5th power of molecular weight increase.[4] In addition to viscosity, a great many other

physical properties of any polymer depend upon the molecular weight.[5]

Molecular Weight Distribution. A given polymer sample is composed of many polymer chains, which in most cases are not of the same length. This variability can be a result of the synthesis process or of possible random scission and cross-linking that can occur upon processing. For economic reasons, it is not possible to separate the various polymer chains by length prior to use; so it is important to characterize this distribution in order to describe the polymer and understand its performance. As with any distribution, no single number is a totally satisfactory descriptor.

The commonly used molecular weight parameters[6] are the number, weight, and z average molecular weight, which are defined, respectively, as:

$$Mn = \Sigma N_i M_i / \Sigma N_i$$

$$Mw = \Sigma N_i M_i^2 / \Sigma N_i M_i$$

$$Mz = \Sigma N_i M_i^3 / \Sigma N_i M_i^2$$

where N_i is the number of moles of species i, and M_i is the molecular weight of species i.

Although there are many different statistical ways to describe any population, the above parameters have been widely used because they are readily understood in physical terms, and they can be measured directly in the laboratory.[7] A fourth parameter, the dispersion index, frequently is used to characterize the breadth of the distribution. This parameter is simply the ratio of the weight to the number average molecular weight, with 1.0 being the lowest possible number (i.e., all chains of exactly the same length). Typical values for commercial polymers are in the 2 to 5 range, with those under 2 considered relatively narrow and those over 2.5 considered broad in distribution. The measurement of these molecular weight averages once was a time-consuming task, but with the development of gel permeation chromatography (GPC), also referred to as size exclusion chromatography, the measurement of these distributions has become commonplace.[8] Units are even available that automatically sample polymerization reactors, process the sample, and perform all necessary calculations to provide data for process control. These units can have multiple detectors, thereby providing compositional distribution as a function of molecular weight.

Branching. The concept of a polymer chain implies two ends per chain. However, because of the nature of the process used to form the polymer, the chain may contain one or more branch points, resulting in multiple ends per chain. These chain ends can have an adverse effect on polymer performance. Branching, molecular weight, and molecular weight distribution have been shown to affect processability as well.[9] The optimum macrostructure often represents a compromise between processing and ultimate performance.

Microstructure

In the formation of elastomers from diolefin monomers such as butadiene or isoprene, there are a number of possible structures. As the control of these structures is critical in obtaining optimum properties, this area has received great attention from the synthesis chemist. The possible isoprene structures are:

For butadiene (no methyl group) the 3,4 form does not exist. The 1,2 addition is referred to as vinyl addition.

The polymer in natural rubber (from the *Hevea brasiliensis* tree) is pure *cis* polyisoprene; gutta percha and balata are composed of the *trans* isomer.

Copolymer Composition

Many of the commercial synthetic elastomers are synthesized from more than one monomer, such as styrene-butadiene and ethylene-propylene rubbers. The properties of the resultant polymer depend on the ratio of the two monomers in the polymer and upon the distribution of the monomers within the chain.

If the monomers are uniformly distributed within the polymer chain, the ratio of monomers will define the flexibility of the polymer chain.[10] Because many properties depend on this chain mobility, polymer composition is carefully controlled. In addition to chain mobility, polymer composition also defines the solubility parameter of the polymer, which is a critical property relative to the type of solvents in which the polymer is soluble, the ability of the polymer to accept and hold oil, and the relative compatibility of the polymer with other polymers.[11] Basically, these properties all involve polymer/solvent interaction, with the difference being the increasing molecular weight of the solvent.

In addition to the relative ratio of the monomers, the arrangement of the units in the chain is important. This arrangement is referred to as the copolymer sequence distribution. In the previous discussion, the assumption was made that the comonomer units were well mixed in the polymer chain. If this is not the case, parts of the chain can reflect properties of the corresponding homopolymer. It is thus possible to produce polymers that have significantly different properties in different parts of the polymer chain. A most dramatic example of this can be found in styrene-butadiene-styrene or styrene-isoprene-styrene thermoplastic elastomers. These triblock polymers behave as cured elastomers at room temperature. The polystyrene blocks have sufficiently different solubility from the polydiene portions that they phase-separate into domains. These polystyrene domains, having multiple chains entering each one, serve as temporary cross-links, in that it is very difficult to pull a styrene block out of the polystyrene domain, and at the same time the domains act as fillers because of their size and hardness. However, once the polymer is heated above the softening point of the domain, the chain can easily be pulled out of the domains, allowing for polymer processing.[12]

In production of most copolymers, the composition incorporated into a polymer chain is generally not the same as that of the monomer ratio, owing to the relative reactivity of different species. As the active chain species can be from either of the two monomers, and either of the monomers can add to any growing end, the following four equations are required to describe the system:

$$-A^* + A \rightarrow -A - A^* \qquad (k_{AA})$$

$$-A^* + B \rightarrow -A - B^* \qquad (k_{AB})$$

$$-B^* + A \rightarrow -B - A^* \qquad (k_{BA})$$

$$-B^* + B \rightarrow -B - B^* \qquad (k_{BB})$$

where the * designates the growing chain end and the k's are the corresponding reaction rate constants. Two relative rate constants, r_A and r_B, called the reactivity ratios, can be used to characterize the nature of the copolymerization, where r_A is defined as k_{AA}/k_{AB} and r_B as k_{BB}/k_{BA}. The copolymer composition will be the same as the monomer mixture only if $r_A = r_B = 1$, which is relatively rare. If r_A and r_B approach zero, the units will tend toward alternation, whereas if $r_A \times r_B = 1$, the copolymer composition is dependent on the monomer composition and will vary with conversion. Tables of reactivity ratios are published, covering many monomer pairs and polymerization conditions.[13] More detailed discussions of copolymerization kinetics can be found in the references.[14,15]

Network Structure

A critical requirement for obtaining engineering properties from a rubbery material is its

existence in a network structure. Charles Goodyear's discovery of vulcanization changed natural rubber from a material that became sticky when hot and brittle when cold into a material that could be used over a wide range of conditions. Basically, he had found a way to chemically connect the individual polymer chains into a three-dimensional network. Chains that previously could flow past one another under stress now had only limited extensibility, which allowed for the support of considerable stress and retraction upon release of the stress. The terms "vulcanization," "rubber cure," and "cross-linking" all refer to the same general phenomenon.

For most rubber systems the network is formed after the polymer is compounded and molded into the desired final shape. Once cross-linked, the material no longer can be processed. If cross-linking occurs prior to compounding or molding, the material is referred to as gelled, and it cannot be used. Most rubber is used in a compounded and cured form.

The one general class of polymers that fall outside this concept is the thermoplastic elastomers, one example of which was discussed previously. These materials can be processed (and reprocessed) at high temperature, yet they maintain properties of cured rubber at use temperatures. This system functions by the formation of either hard plastic, crystalline, or ionic domains that, at use temperature, act as cross-link sites because multiple chains are involved in the domains. Upon heating, the integrity of these domains breaks down, and the polymer chains can easily flow past one another. It should be noted that at use temperatures these systems have a three-dimensional network. Such systems tend to show more creep and stress relaxation than cured systems, as the network is formed via weaker secondary effects rather than primary chemical bonds. These problems become more severe as the use temperature is increased because ultimately the network cannot remain intact at processing temperatures. For any network, its structure is important in defining the performance of the final product. Such properties as modulus, hardness, and tensile strength are dependent on the network density.[16]

RUBBER PROPERTIES

Glass Transition Temperature

In order for a polymer to behave as a rubbery material, it is necessary for the chain to have great mobility. As the temperature is lowered, the ability of the chain segments to move decreases until a temperature is reached where any large-scale motion is prevented. This temperature is referred to as the glass transition temperature. Below this temperature the rubber becomes a glassy material—hard and brittle; above this temperature amorphous plastics, such as polystyrene, can exhibit rubbery character if the molecular weight is sufficiently high. All rubbery materials, then, must have glass transition temperatures below room temperature. For good low-temperature properties, it obviously follows that a low glass transition temperature polymer is required. The control of the glass transition temperature of the polymer is critical for many properties other than low temperature use. For example, the wet traction and wear of a passenger tire have been shown to depend directly on the polymer glass transition temperature.[17]

Control of this property is possible by controlling the structure of the polymer chain. Monomers with bulky side groups restrict chain mobility and thus raise the glass transition temperature. The composition of copolymers and the ratio of polymer blends often are determined by the desired glass transition temperature of the final product.

Crystallinity

Polymer chains can exist in spatial arrangements that are orderly enough to allow the chains to form crystalline structures. The existence of strong interchain interactions via polar forces, hydrogen bonding, or ionic groups can facilitate crystallization. The existence of crystallization is very important

for many plastics and fibers, but crystallinity cannot be appreciably present in rubbery materials, as the corresponding restriction in chain mobility could preclude the very chain mobility needed for rubbery properties. Rubbery materials must have both their melting temperature (if they have a crystalline point) and their glass transition temperature below the use temperature.

Crystallinity can be reduced by disruption of the order in the chain by copolymerization.[18] For example, both polyethylene and polypropylene are crystalline plastics, whereas ethylene/propylene rubber produced at about a 50/50 ratio is an amorphous elastomer. Compositional excursions much outside this range lead to crystalline materials.[19] For some materials, such as natural rubber, that are close to crystallizing, stretching the chains can align them sufficiently for crystallization to occur. Such polymers can exhibit excellent gum properties and improved strength in the uncured state that greatly facilitate processing. Attempts have been made with some success to produce other polymers that exhibit this property of natural rubber. Although the melting temperature can be matched by appropriately disrupting the crystallizable structure through controlled introduction of another monomer, an exact match is not possible because the extent of crystallinity and the kinetics of crystallization will differ.

RUBBER USE

Compounding

The rubber industry began when Charles Goodyear developed the first useful rubber compound: natural rubber plus sulfur. The concept of mixing materials into rubber to improve performance is still of primary importance today. Without compounding, few rubbers would be of any commercial value. Any given rubber application will have a long list of necessary criteria in addition to cost, encompassing appearance, processing, mechanical, electrical, chemical, and thermal properties. Developing such compounds requires a broad knowledge of material science and chemistry combined with experience. The use of designed experiments (see Chapter 4) can greatly facilitate selecting the optimum compound formulation.

The major components in a compound are curatives, reinforcing agents, fillers, plasticizers, and antidegradants.

Curatives. The function of curatives is to cross-link the polymer chains into a network; the most common ones are the sulfur type for unsaturated rubber and peroxides for saturated polymers. Chemicals called accelerators may be added to control the cure rate in the sulfur system; these materials generally are complex organic chemicals containing sulfur and nitrogen atoms. Stearic acid and zinc oxide usually are added to activate these accelerators. Metal oxides are used to cure halogenated polymers such as polychloroprene or chlorosulfonated polyethylene.

Reinforcing Agents. Carbon black and silica are the most common reinforcing agents. These materials improve properties such as tensile strength and tear strength; also, they increase hardness, stiffness, and density and reduce cost. Almost all rubbers require reinforcement to obtain acceptable use properties. The size of the particles, how they may be interconnected (structure), and the chemical activity of the surface are all critical properties for reinforcing agents.

Fillers. Fillers are added to reduce cost, increase hardness, and color the compound. They generally do not provide the dramatic improvement in properties seen with reinforcing agents, but they may have some reinforcing capability. Typical fillers are clays, calcium carbonate, and titanium dioxide.

Plasticizers. These materials are added to reduce the hardness of the compound and can reduce the viscosity of the uncured compound to facilitate processes such as mixing and extruding. The most common materials are petroleum-based oils, esters, and fatty acids. Critical properties of these materials are their compatibility with the rubber and their

viscosity. Failure to obtain sufficient compatibility will cause the plasticizer to diffuse out of the compound. The oils are classified as aromatic, naphthenic, or paraffinic according to their components. Aromatic oils will be more compatible with styrene-butadiene rubber than paraffinic oils, whereas the inverse will be true for butyl rubber. The aromatic oils are dark colored and thus cannot be used where color is critical, as in the white sidewall of a tire. The naphthenic and paraffinic oils can be water-white and are referred to as nonstaining.

Antidegradents. This group of chemicals is added to prevent undesirable chemical reactions with the polymer network. The most important are the antioxidants, which trap free radicals and prevent chain scission and cross-linking. Antiozonants are added to prevent ozone attack on the rubber, which can lead to the formation and growth of cracks. Antiozonants function by diffusion of the material to the surface of the rubber, thereby providing a protective film. Certain antioxidants have this characteristic, and waxes also are used for this purpose.

Processing

A wide range of processes are used to convert a bale of rubber into a rubber product such as a tire. The first process generally will be compounding. Typical compounding ingredients were discussed previously. In many compounds more than one rubber may be needed to obtain the performance required. Uncured rubber can be considered as a very high-viscosity liquid; it really is a viscoelastic material possessing both liquid and elastic properties. Mixing materials into rubber requires high shear, and the simplest method is a double roll mill in which the rubber is shear-mixed along with the other compounding ingredients in the bite of the mill. Large-scale mixing is most commonly done with a high-shear internal mixer called a Banbury. This mixing is a batch process although continuous internal mixers also are used.

The compounded rubber stock will be further processed for use. The process could be as simple as injection- or transfer-molding the stock into a hot mold where it is cured. Extrusion of shaped stock frequently is used to make hose or tire tread material. Another common process is calendering, in which a fabric is passed through rolls where rubber stock is squeezed into the fabric, to make rubber fabric for roofing membranes or body plies containing the cord for tire construction. The actual construction of the final product can be very complex. For example, a tire is built of many different rubber stocks and components. Different compounds are used for sidewalls, treads, body plies, and bead areas. Steel or fiber cords are required for the tire body, and steel is required to provide the strength needed in the bead area, which holds the tire to the rim. All of these components must be assembled with high precision so that the final cured unit can operate smoothly at high speeds. Tires must perform at hot and cold temperatures under wet and dry conditions for 30,000 or more miles. Although tires appear to be simple round black objects, they obviously are complex high-technology composite structures.[20–22]

NATURAL RUBBER

More than 500 years ago, the people of Central and South America were using a product that they collected from certain trees to make balls and to coat fabric to make it waterproof. This material they called *cauchuc*, which means "weeping wood." Today we know the tree as the *Hevea brasiliensis* and the material as natural rubber. Although a number of plants produce rubber, the only significant commercial source is the *Hevea* tree. Natural rubber had only limited applications until 1839, when Charles Goodyear found that when combined with sulfur and heated, the material changed into cured rubber with properties much as we know them today. The development of the pneumatic tire in 1845 combined with the dramatic growth of the automotive industry led to a rapid increase in the demand for natural rubber.

Although the tree was indigenous to Brazil, seeds were taken to England where they were germinated, and the plants were sent to the Far East. Rubber plantations were in existence by the late 1800s, and in the 1920s plantations were begun in West Africa. Because of a leaf blight disease, essentially all natural rubber now comes from plantations in Africa and the Far East rather than Central and South America.[21]

The production process starts with the trees. Over the years considerable biological research has been done to produce trees that grow faster, produce more latex, and are resistant to wind and disease damage. Once such an improved tree has been identified, buds are grafted from the tree onto root stock. All such trees are referred to as clones and will have the same characteristics. It typically takes 6 to 7 years of growth before a tree is ready for rubber recovery. Peak rubber production is reached at 12 to 15 years of age. Another major development in improving tree performance has been the use of tree stimulants, which have resulted in an overall yield increase of 30 percent without adverse effects on the trees.

The production process of natural rubber in the tree is not yet fully understood. However, it involves a long series of complex biochemical reactions that do not involve isoprene as a monomer, even though the resulting polymer is 100 percent *cis* 1,4 polyisoprene. Because the tree makes the product, the rubber production process is really one of recovery.

The recovery process starts with tapping of the tree, which involves manually removing thin sections of bark at an angle so that as the latex is exuded from the damaged living latex cells, it will flow down to be collected in a cup. The depth of the cut is critical, as a too-shallow cut will not allow optimum latex yield, and too deep a cut will damage the tree. The damaged latex vessels will seal off after several hours. A preservative is added to the collection cup to prevent coagulation of the latex. The tapper then collects the latex and takes it to a collection station, from which it is shipped to the rubber factory. Additional preservative is added at the collection station to assure stable latex. The latex contains around 30 to 45 percent rubber. Approximately 3 percent of the solid material is nonrubber materials, consisting primarily of proteins, resins, and sugars.

The latex is processed in one of several ways, depending on the desired final product. If it is going to be used in the final product application in the latex form (such as for dipped goods and adhesives), it will be concentrated to 60 percent or higher solids. The most common process is centrifugation. The latex separates into the high-solids product and a low-solids skim material that contains a much higher percentage of the nonrubbery components; rubber produced from skim is generally of a lower quality because of these impurities. A variety of chemicals can be added to the latex to provide the necessary preservation and mechanical stability. As the final use of the latex will involve destabilizing the rubber particles, care must be taken not to overstabilize the latex concentrate. Being a natural product, the latex tends to change upon aging, a factor that also must be compensated for in the process.

Dry rubber is produced from the latex first by dilution, then by coagulation with organic acids, and finally by formation into sheets or crepe. Rubber smoked sheets are made by working the coagulated sheets between rolls to remove as much of the nonrubbery components as possible, followed by drying for up to a week in a smokehouse. The smoke serves as a fungicide that prevents biological attack of the rubber. To provide a more well-defined uniform product, Technically Specified Rubber (TSR) processes have been developed. These processes involve converting the coagulated rubber into rubber crumb, which is further washed, dried, and baled. Constant-viscosity grades of natural rubber have been developed by chemically reacting the aldehyde groups, which otherwise would lead to cross-linking upon storage. In the pale crepe process the latex is carefully selected for colored bodies (from carotene) and treated with sodium bisulfite to stop enzyme activity. The rubber is extensively washed through

rollers to remove serum, as this can lead to yellowing. The wet crepe is dried under carefully controlled temperatures and in the absence of light to assure optimum properties. The previously described processes involve considerable investment and are practical only for plantation operations. Small shareholders may allow their latex to coagulate naturally and sell it to processors. Such a product will vary greatly, depending on the specific history of each rubber slab.[23]

Uses

The largest use of natural rubber is in the manufacture of tires. Over 70 percent of its consumption is in this area. The next largest use is as latex in dipped goods, adhesives, rubber thread, and foam. These uses account for approximately another 10 percent. The remainder is used in a variety of applications such as conveyor belts, hoses, gaskets, footwear, and antivibration devices such as engine mounts.

Because of the high stereoregularity of natural rubber, the units in the polymer chain can form very orderly arrangements, which result in crystallization upon storage at low temperatures or upon stretching. Stored crystallized rubber may be converted to its original amorphous state by heating. Several crystalline forms are reported, with melting points varying from 14°C to 36°C.[24] Although crystallization upon storage can be a problem to users, the ability to crystallize reversibly upon stretching accounts for many of the unique properties of natural rubber. Specifically, the ability of natural rubber to be used as a gum polymer (unfilled vulcanizate) depends on this property. The crystallites that form act both as filler and as temporary cross-links, providing high tensile properties.[24] In tire fabrication this property is reflected in natural rubber stocks possessing high green (uncured) strength and excellent building tack. In the final product, strain-induced crystallization provides tear and cracking (cut growth) resistance.

Although many other polymers can crystallize, only natural rubber has been found to have the necessary combination of rate of crystallization, degree of crystallization, and melting point to provide all of the properties discussed above. Natural rubber has good flexibility and resilience. In truck and bus tires, it is used extensively with blends of polybutadiene to give the low heat buildup needed as well as wear and cut growth resistance. In passenger tires, natural rubber is used in the sidewalls and carcass areas. These areas require the building tack, ply adhesion, and hot strength properties that it imparts. Synthetic rubber is used almost exclusively in passenger tire treads.

The largest use of natural rubber latex is in the dipped goods area. Products include balloons, surgical and examination gloves, and prophylactics. The rapid spread of AIDS has led to a dramatic increase in the surgical and examination glove market. Latex also is used to make adhesives, rubber thread, and foams. However, natural rubber is being replaced in these two latex areas by urethanes and synthetic rubber latex. Although natural rubber latex is a mature product, research continues on improving its uniformity, stability, and performance.[26]

POLYISOPRENE

Faraday discovered in 1826 that natural rubber was composed of a hydrocarbon with a ratio of five carbons to eight hydrogen atoms; and in 1860, G. Williams isolated isoprene by collecting the distillate from the heating of natural rubber. By 1887, scientists in France, England, and Germany had converted isoprene back into a rubbery material. Because this offered a potential for manufactured "natural" rubber, research was undertaken to find ways to obtain isoprene from sources other than rubber itself.[27]

Monomer Production

The primary source of isoprene today is as a by-product in the production of ethylene via naphtha cracking. A solvent extraction process

very similar to that used for butadiene is employed. Much less isoprene is produced in the crackers than butadiene, so the availability of isoprene is much more limited. Isoprene also may be produced by the catalytic dehydrogenation of amylenes, which are available in C-5 refinery streams:

It also can be produced from propylene by a dimerization process, followed by isomerization and steam cracking. A third route involves the use of acetone and acetylene, produced from coal via calcium carbide. The resulting 2 methyl butyn-3-ol-2 is hydrogenated to methyl butanol and subsequently dehydrogenated to give isoprene. The plants that were built on these last two processes have been shut down, evidently because of the relatively low cost of the extraction route.

Polymer Production Process

The free-radical catalysts were found to produce a product that did not have the tack, green strength, or gum tensile of natural rubber. Whereas natural rubber is an essentially pure *cis*-1,4 structure, the emulsion product was of mixed microstructure. This precluded the ability of the latter to undergo stress-induced crystallization, which is required to obtain many of the desired natural rubber properties.

In 1955 investigators from the Firestone Tire and Rubber Company and the B. F. Goodrich Company announced the synthesis of polyisoprene with over 90 percent *cis*-1,4 structure. The work at Firestone was based on lithium metal catalysts, whereas the work at Goodrich was the result of using Ziegler-Natta type coordination catalysts.[28,29]

Use

Although considerable interest was generated by these discoveries, their commercial success has been rather limited. The lithium-based polymers were found to produce up to 94 percent *cis*, which still was not high enough to provide the properties of natural rubber. Polymers made with the coordination catalysts have *cis* contents of up to 98 percent, providing products that can more closely serve as replacements for natural rubber than the lithium-based polymers. In comparison with natural rubber, they offer the advantage of very high rubber content and excellent uniformity in quality. For economic reasons, polyisoprene has seen only limited success. Several of the plants built to produce polyisoprene have been either shut down or converted for use to produce other polymers. In terms of synthetic rubber production in 1988, only 129,000 metric tons were produced (excluding centrally planned economy countries), which represented only 3 percent of total synthetic rubber production. Since that time an additional 45,000-metric-ton plant has been closed. In comparison, in 1989 natural rubber usage was 4,114,000 metric tons.[30] Evidently because of its strategic importance, the former Soviet Union continues to rely heavily on polyisoprene, with a 1990 capacity of 960,000 metric tons.[31]

STYRENE-BUTADIENE RUBBER

The largest-volume synthetic rubber consumed is styrene-butadiene rubber (SBR). In 1990, SBR solid rubber accounted for 35 percent of all synthetic rubber. If SBR latex and carboxylated SBR latex are included, its share increases to 55 percent. The major application of solid SBR is in the automotive and tire industry, accounting for approximately 70 percent of the use. Therefore, SBR has been tightly tied to the tire business.[32]

Initially, SBR was developed as a general purpose alternate material to natural rubber. In the United States the thrust came early in World War II when the U.S. supply of natural rubber was cut off. The basic technology was

developed in Germany in the late 1920s, and by 1939 Germany had 175,000 metric tons of capacity in place. The first U.S. production was 230 metric tons in 1941, but by 1945 more than 850,000 metric tons of capacity were in place. Basically, in a period of five years the emulsion SBR business as we know it today was put in place. By 1973 U.S. capacity had increased to almost 1,400,000 metric tons, but in 1989 it had contracted to 881,000 tons—close to the 1945 capacity.[33]

Monomer Production

The production of butadiene monomer is discussed below in the polybutadiene section. Styrene is producted by two processes: The largest volume process involves the alkylation of benzene with ethylene to give ethyl benzene, which is then dehydrogenated to give styrene.[34]

$$\underset{\text{Benzene}}{C_6H_6} + \underset{\text{Ethylene}}{C_2H_4} \longrightarrow \underset{\text{Ethyl benzene}}{C_6H_5CH_2CH_3}$$

$$C_6H_5CH_2CH_3 \longrightarrow \underset{\text{Styrene}}{C_6H_5CHCH_2} + H_2$$

The second process involves the coproduction of propylene oxide.[35] Ethylbenzene is oxidized to its hydroperoxide, which is reacted with propylene to give the coproduct propylene-oxide and methyl phenyl carbinol, which is then dehydrated to styrene.

Polymer Production Process

SBR is produced by two different processes, emulsion and solution. The emulsion process involves a free radical mechanism, whereas solution SBR is based on alkyllithium catalysis.

Emulsion Process

The formula developed to provide SBR during World War II was standardized, with all rubber plants owned by the U.S. government. The standard recipe is listed below:[36]

Component	*Parts by Weight*
Butadiene	75
Styrene	25
n-Dodecyl mercaptan	0.5
Potassium peroxydisulfate	0.3
Soap flakes	5.0
Water	180

Initiation occurs through reaction of the persulfate with the mercaptan, as shown below:

$$\underset{\text{Potassium persulfate}}{K_2S_2O_8} + \underset{\text{Alkylmercaptan}}{2RSH} \longrightarrow \underset{\text{Mercaptyl free radical}}{2RS\cdot} + KHSO_4$$

$$RS\cdot + \underset{\text{monomer}}{M} \longrightarrow RSM\cdot$$

Chain propagation occurs by the growing-chain free radical attacking either butadiene or styrene monomer. The active radical chain can react with mercaptan to form a new

$$C_6H_5CH_2CH_3 + O_2 \longrightarrow \underset{\text{Ethylbenzene hydroperoxide}}{C_6H_5CHOOHCH_3}$$

$$C_6H_5CHOOHCH_3 + \underset{\text{Propylene}}{CH_3CHCH_2} \longrightarrow \underset{\text{Propylene oxide}}{CH_3\underset{\diagdown O \diagup}{CH}CH_3} + \underset{\text{Methyl phenyl carbinol}}{C_6H_5CHOHCH_3}$$

$$C_6H_5CHOHCH_3 \longrightarrow \underset{\text{Styrene}}{C_6H_5CHCH_2} + H_2O$$

mercaptyl radical and a terminated chain. The mercaptyl radical then can initiate an additional chain. The molecular weight of the chain P can be controlled by the concentration of mercaptan via this chain transfer mechanism.

$$P\cdot + RSH \rightarrow PH + RS\cdot$$

$$RS\cdot + M \rightarrow RSM\cdot$$

Termination also can occur by the reaction of two free radicals, through either combination or disproportionation reactions.

$$P\cdot + P\cdot \rightarrow P\text{-}P \quad \text{Combination (bimolecular coupling)}$$

$$P\cdot + P\cdot \rightarrow P\text{-}CH = CH_2 + PH \quad \text{Disproportionation (hydrogen free radical transfer)}$$

The mercaptyl radical also can react with growing chains, to lead to termination.[37]

Polymerization is carried out at 50°C until conversion of 70 to 75 percent is reached, at which time the polymerization is terminated by the addition of a free radical scavenger such as hydroquinone. Polymerization beyond this point results in excessive free radical attack on the polymer chains. Products made under such conditions have poor properties due to excessive branching and gelation. Unreacted butadiene and styrene are removed by flashing and steam stripping. Antioxidant is added to the latex, followed by coagulation with the addition of polyelectrolytes and salt-acid. The coagulated crumb then is washed, dried, baled, wrapped, and packaged for shipment. Because of the soap and other chemicals in the formulation, most emulsion polymers will contain about 7 percent of nonrubber residues. The emulsion process flow sheet is shown in Fig. 18.1.

It soon was discovered that polymers made at lower temperatures had significantly better properties, especially tire threads, because they contained less low molecular weight material. Attempts simply to lower the polymerization temperature were unsuccessful because of the resulting low polymerization rate. Success finally was achieved through the use of a "redox" system. This system involves an oxidizing agent such as peroxides or hydroperoxides and a reducing agent such as water-soluble transition metal salts. These systems are very active, even at temperatures as low as 0°C.

Mercaptans again are used as chain transfer agents to provide a mechanism for molecular weight control. Commercially these types of polymerization are carried out at 5°C and are referred to as "cold" polymerizations to differentiate them from the previously discussed "hot" systems. A typical formula is listed below.[38]

Component	*Parts by Weight*
Butadiene	71
Styrene	29
tert-Dodecyl mercaptan	0.18
p-Methane hydroperoxide	0.08
Ferrous sulfate heptahydrate	0.03
Trisodium phosphate decahydrate	0.50
Tetrasodium ethylenediaminetetraacetate	0.035
Sodium formaldehyde sulfoxylate	0.08
Rosin acid soap	4.5
Water	200

The improved wear properties of cold polymerization SBR led to the very rapid replacement of hot SBR for most applications. This change was relatively easy to make, as all of the equipment could be used with the only modification required being the addition of reactor cooling, which is achieved with either the reactor jacket, internal coils, or both.

It later was found that even more improvements could be realized by polymerizing to very high molecular weights and then adding petroleum-based oils to the latex prior to coagulation. The oil is absorbed by the rubber, which upon coagulation produces oil-extended polymers. For tread applications, oils of higher aromatic content were preferred because of their excellent compatibility with the rubber. Typically 37.5 parts of oil are added although grades containing up to 50 parts have been produced. Very high molecular weight polymers thus can be processed

Fig. 18.1. The production of styrene-butadiene rubber. (*Reprinted with permission from* The Vanderbilt Rubber Handbook, *13th ed., R. F. Ohm, Ed. Copyright 1990, the R. T. Vanderbilt Company, Inc.*)

without requiring excessive energy to mix them and avoiding polymer degradation. Carbon black masterbatches also are produced. In this process carbon black is added to the latex prior to coagulation, and the black, along with oil, is incorporated into the latex in the coagulation step. These products offer the user the advantage of not having to handle free black in their mixing operation, and can provide additional compounding volume for manufacturers with limited mixing capacity.

Not all emulsion SBR is converted to dry rubber for use. There are a variety of applications where the latex can be used directly in the final fabrication process. This

technology logically grew out of the latex technology developed for natural rubber. For latex applications the particle size distribution can be critical because of its impact on viscosity as well as possible performance variables, as when used to provide impact strength in plastics such as ABS.[39] Careful control of the mechanical stability of the latex also is critical, as these systems must destabilize under relatively mild conditions such as those in a coating operation. A number of processes have been developed to control particle size via partial destabilization of the latex. Among the commercial methods are: careful control of a freeze–thaw cycle, controlled shear agitators, high pressure colloid mills, and the addition of chemicals such as hydrocarbons or glycols.

A special variation of SBR latex containing terpolymerized vinyl pyridine is used in the tire industry to provide adhesion of organic fiber tire cords to rubber stock. The vinyl pyridine SBR latex is combined with resins and coated on the fiber by a dipping process, followed by a controlled temperature and tension process to set the adhesive as well as to control the shrinkage properties of the cord.

Solution Process

The discovery of the ability of lithium-based catalysts to polymerize isoprene to give a high *cis* 1,4 polyisoprene was rapidly followed by the development of alkyllithium-based polybutadiene. The first commercial plant was built by the Firestone Tire and Rubber Company in 1960. Within a few years the technology was expanded to butadiene/styrene copolymers, with commercial production under way toward the end of the 1960s.

The copolymerization with alkyllithium to produce uniformly random copolymers is more complex for the solution process than for emulsion because the reactivity ratios differ greatly. For nonpolar solvent systems values for r_B of 14 and r_S of .07 are reported.[39] Although the product of the reactivity ratios is close to one and the process thus is random from a statistical basis, the actual composition is not uniform in the polymer chain. Because of the extremely high rate of reaction of the styryllithium anion with butadiene, the polymerization very heavily favors the incorporation of butadiene units as long as reasonable concentrations of butadiene are present. This observation initially was somewhat confusing because the homopolymerization rate of styrene is seven times that for butadiene. However, the cross-propagation rate is orders of magnitude faster than either, and it therefore dominates the system. For a 30 mole percent styrene charge the initial polymer will be almost pure butadiene until most of the butadiene is polymerized. Typically two-thirds of the styrene charged will be found as a block of polystyrene at the tail end of the polymer chain.

Several methods have been proposed to overcome this problem. In one, the styrene and part of the butadiene are charged initially with butadiene metered at a rate equivalent to its incorporation into the chain. A second approach involves adding both monomers at a relatively slow rate so that the equilibrium monomer concentration reaches a pseudo-steady state that will produce polymer at the desired composition.[40] This process can be done in either a batch or a continuous mode.[41]

In addition to these reaction engineering approaches to produce uniformly random copolymer, the chemistry may be changed by the addition of polar agents such as amines and ethers. This action results in bringing the reactivities much closer together. The change in chemistry also is reflected in the microstructure of the butadiene portion of the polymer. Whereas in the nonpolar system the vinyl content of the butadiene portion is around 10 percent, in polar systems vinyl contents of 30 to 40 percent typically are obtained when a 20 percent styrene polymer is randomized. Higher styrene contents require higher modifier levels, resulting in even higher vinyl contents. An added complication with polar modifiers is their ability to react with the growing chain, resulting in undesired termination.[42]

In spite of these complications, all recent U.S. expansions or announced plants for SBR have been for solution polymers. This fact is

even more surprising, considering the significant excess emulsion SBR capacity in existence.

The driving force toward solution SBR is its improved properties for passenger tire treads. As the polymerization recipe consists of monomer, solvent, possibly trace polar agents, and trace initiator, the polymer contains essentially all-rubber hydrocarbon. The only major nonrubber constituent will be the antioxidant added to all unsaturated polymers. In addition to the high rubber hydrocarbon content, solution polymers can offer more linear chains and narrower molecular weight distributions than emulsion polymers.[43] Advantage also can be taken of the relatively stable growing chain end to chemically modify the polymer to improve its interaction with carbon black in tire compounds. This modification can lead to a dramatic reduction in rolling resistance, which is critical for automotive manufacturers who must meet government-mandated fuel-economy targets. One study comparing an emulsion polymer, its solution counterpart, and a chemically modified version of the solution polymer showed a 23 percent hysteresis reduction in going from emulsion to solution and an additional 15 percent reduction for the chemically modified polymer, to provide an overall reduction of 38 percent.[44] To date, this type of control has not been demonstrated for the emulsion technology.

POLYBUTADIENE

Next to SBR, polybutadiene is the largest-volume synthetic rubber produced. Consumption was approximately 1,140,000 metric tons in 1990 (excluding centrally planned economy countries).[45]

Monomer Production

Butadiene monomer can be produced by a number of different processes. The dominant method of production is as a by-product from the steam cracking of naphtha to produce ethylene. The butadiene is recovered from the C-4 fractions by extractive distillation.[46] "On-purpose" butadiene is generally produced by dehydrogenation or oxidative dehydrogenation of four-carbon hydrocarbons.[47,48]

Polymer Production Process

Polybutadiene is produced by free-radical emulsion, alkali metal solution, and transition metal coordination solution processes. Most production is based on the solution processes because of the ability to obtain preferred microstructures by these routes.

Alkali-metal-based polymer (usually organolithium) produces a product with about 36 percent *cis*, 54 percent *trans*, and 10 percent vinyl. The polymerization process is conducted in an aliphatic hydrocarbon under an inert atmosphere in either a batch or a continuous mode. Because of the characteristics of this polymerization system, polymers of extremely narrow molecular weight distribution and low gel can be produced.[49] The narrowest distribution is produced via batch polymerization.

Upon the addition of polar agents, such as ethers or amines, the organolithium initiators can produce polybutadienes with vinyl contents up to 100 percent.[50] The vinyl content can be controlled by the ratio of modifier to catalyst and the polymerization temperature, with lower temperatures favoring increased vinyl formation. Even at the high vinyl contents such polymers do not crystallize, because of the atactic nature of the vinyl units.

High *cis* polybutadiene is produced via solution processes using Ziegler-Natta type transition metal catalysts. The major commercial catalysts of this type are based on titanium, cobalt, nickel, and neodymium.[51] Typically the transition metal is used in the form of an organometallic compound, which is reacted with an organoaluminum or organoaluminum halide as a reducing agent to give the active species at the desired transition state. The cobalt catalysts appear to require aromatic solvents, whereas aliphatic hydrocarbons can be used for the others. Because of the active nature of transition metals, the

polymer solutions are treated to deactivate or remove such materials from the final product. All of these catalysts produce products with 95 percent or higher *cis* content. The neodymium system is reported to produce the highest *cis* (98–99%).[52]

All of the solution processes require high efficiency in recovering the solvent. The most widely used process consists of termination of the polymerization and the addition of antioxidant to the polymer solution. The solution may be treated to remove catalyst residue and then transferred into an agitated steam stripping vessel in which unreacted monomer and solvent are flashed off, leaving the rubber as a crumb slurry in water. The water/crumb slurry then is dewatered and dried. The recovered monomer/solvent is recirculated to a series of distillation columns to recover monomer and purify the solvent. As both the anionic and the coordination catalyst systems are highly sensitive to impurities such as water, the purification system is very critical for satisfactory process control. (Fig. 18.2).

Oil-extended polybutadiene also is produced. The production process is basically the same with the exception that a higher molecular weight polymer is produced, which upon oil extension has about the same bulk viscosity as the nonextended polymers. The oil is added to the polymer solution prior to the desolventization step. Oil carbon black masterbatches also are produced. In these processes the carbon black must be ground or debeaded in some mechanical manner and then transferred as a slurry, either in solvent or water, to the polymer solution.

The emulsion polybutadiene process is similar to that described for emulsion SBR.

Fig. 18.2. Flow diagram for a typical solution process for manufacture of polybutadiene. (*Courtesy Firestone Synthetic Rubber and Latex Co.*)

Uses

The major use for polybutadiene is in tires, with over 70 percent of the polymer produced used by the tire industry. Cured polybutadiene has excellent low-temperature properties, high resiliency, and good abrasion resistance due to its low glass transition temperature. However, this same fundamental property also leads to very poor wet skid resistance. For this reason, polybutadiene is blended with other polymers such as natural rubber and SBR for use in tread compounds. In general, polybutadiene is a poorer-processing polymer than SBR, but this is generally not a problem as it is blended with other polymers in use. The very high *cis* polymers have the potential for strain-induced crystallization, which can lead to improved green strength and increased cut growth resistance in the cured product. High *cis* polybutadiene is reported to have a melting point of 6°C.[53]

The other major use for polybutadiene is as an impact modifier in plastics, in particular high impact polystyrene (HIPS) and acryolonitrile-butadiene-styrene resin (ABS). In the HIPS application the rubber is dissolved in the styrene monomer, which is then polymerized via a free radical mechanism. A complex series of phase changes occurs, resulting in small rubber particles containing even smaller polystyrene particles being incorporated into a polystyrene matrix. The rubber is added to impart impact strength. Because of the unique morphology that is formed, low levels of rubber (typically around 7%) provide rubbery particles having a volume fraction of 30 to 40 percent. This morphology leads to high impact at very low rubber levels, providing for good stiffness and hardness.[54]

Polybutadiene may be used in a similar manner in ABS production via a mass process. However, most is emulsion-polymerized for use as a latex in the emulsion ABS process.

ETHYLENE/PROPYLENE RUBBER

There are two general types of polymers based on ethylene and propylene: ethylene/propylene rubber (EPM) and ethylene/propylene terpolymer (EPDM). EPM accounts for approximately 20 percent of the polyolefin rubber produced. Comprising a totally saturated polymer, these materials require free radical sources to cross-link. EPDM was developed to overcome this cure limitation. For EPDM a small amount (less than 15%) of a nonconjugated diene is terpolymerized into the polymer. One of the olefinic groups is incorporated into the chain, leaving its other unsaturated site free for vulcanization. This ensures that the polymer backbone remains saturated, with corresponding stability, while still providing the reactive side group necessary for conventional cure systems. The nonconjugated dienes used commercially are ethylidene norbornene, 1,4 hexadiene, and dicyclopentadiene. The selection of the termonomer is made on the basis of the reactivity of the termonomer, both in polymerization and in vulcanization. The estimated 1990 worldwide consumption (excluding CPEC) was 593,000 metric tons.

Monomer Production

Ethylene and propylene are produced primarily by the cracking of naphtha. They also are available from the fractionation of natural gas.

Ethylidene norbornene is produced by reacting butadiene with cyclopentadiene. 1,4 Hexadiene is produced from butadiene and ethylene:

$$CH_2{=}CH{-}CH{=}CH_2 + CH_2{=}CH_2 \longrightarrow \underset{\text{1,4 Hexadiene}}{CH_2{=}CH{-}CH_2{-}CH{=}CHCH_3}$$

Dicyclopentadiene is obtained as a byproduct from the cracking of heavy feedstocks to produce ethylene.

Polymer Production

There are two processes used to produce EPM/EPDM: solution and suspension. In either case a Ziegler-Natta type catalyst is

used (aluminum alkyl or aluminum alkyl chlorides and a transition metal salt). The most generally used transition metal is vanadium in the form of the tetrachloride or the oxytrichloride.[55] The solution process is similar to that used for other solution polymers. The polymer cement can be finished by stream stripping and drying of the resulting crumb.[56]

In the suspension process, the polymer is suspended in the monomer propylene. This process offers the advantages of being able to operate at higher solids (owing to the lower viscosity of a suspension than that of a solution at comparable solids), simple heat removal by the evaporative cooling of the propylene, more uniform reactor temperature profile, and ease of production of high molecular weight or semicrystalline polymers, as a polymer solution is not required.[57]

Announcement recently was made of the use of a specially developed titanium-based catalyst in the suspension process for EPM and EPDM where the termonomer is low-boiling. The advantages claimed, in addition to those characteristic of the suspension process, are better structural control and high catalyst efficiency, resulting in a high-purity product without requiring catalyst removal.[58]

The polymer composition for both EPM and EPDM is usually in the 40/60 to 60/40 ethylene/propylene ratio. Outside these ranges, the polymer will start to crystallize because of either polyethylene or polypropylene blocks.

Use

EPM/EPDM polymers exhibit outstanding resistance to heat, ozone, oxidation, weathering, and aging due to the saturated backbone. They have low density, accept aliphatic and naphthenic oils well, and maintain acceptable properties at high filler loadings. They are used in single-ply roofing, wire and cable installation, automotive parts, impact modification of polypropylene, and viscosity index additives for automotive oils and in producing thermoplastic olefin elastomers by blending with polypropylene, which may be partially grafted or cross-linked by dynamic vulcanization. Although at one time EPDM was expected to become the major polymer for tires, this market has not materialized for a variety of processing and performance reasons.[59]

BUTYL RUBBER

Butyl rubber is one of the older synthetic rubbers, having been developed in 1937. Because of the saturated nature of a polyolefin elastomer, the commercial polymer is actually a copolymer of isobutylene and isoprene. The isoprene is added to provide cure sites. In addition, halogenated (bromo or chloro) derivatives are available. The halogenated products improve the mixing and cure compatibility with the more common unsaturated rubbers such as natural or styrene-butadiene rubber.

Monomer Production

Isobutylene is obtained as a by-product from petroleum and natural gas plants. The monomer must be highly purified to assure high molecular weight.

Production Process

Butyl rubber is produced at very low temperature (below $-90°C$) to control the rapid exotherm and to provide high molecular weight. The process consists of charging isobutylene along with isoprene (2–4%) with an inert diluent such as methyl chloride to a reactor to which a Friedel-Crafts catalyst is added. The polymerization is very rapid, and the polymer forms in a crumb or slurry in the diluent. Heat is removed via a jacket by evaporating materials such as ethylene. The slurry is steam-stripped to remove all volatiles. The catalyst is neutralized, and antioxidants are added to the slurry prior to drying.[60,61]

The halogenated derivatives are produced by the direct addition of the halogen to a solution of the isobutylene-isoprene polymer.

Properties and Use

The most important characteristic of butyl rubber is its low permeability to air. This property accounts for its major use in inner tubes and tire inner liners. Because of the poor compatibility of butyl with other rubbers (with respect to both solubility and cure), the halobutyls are preferred.

Other uses for butyl rubber are automotive mechanical parts (due to the high damping characteristics of butyl), adhesives, mastics, and sealants.[62]

NITRILE RUBBER

Nitrile rubber was invented at about the same time as SBR in the German program to find substitutes for natural rubber.[63] These rubbers are copolymers of acrylonitrile/butadiene, containing from 15 to 40 percent acrylonitrile. The major applications for this material are in areas requiring oil and solvent resistance. The estimated worldwide consumption in 1990 (excluding CPEC) was 252,000 metric tons.[64]

Monomer Production

The production of butadiene is discussed in the polybutadiene section. Although several routes have been developed to produce acrylonitrile, almost all now is produced by the catalytic fluidized-bed ammoxidation of propylene:

$$\underset{\text{Propylene}}{CH_2{=}CHCH_3} + \underset{\text{Ammonia}}{NH_3} + O_2 \rightarrow \underset{\text{Acrylonitrile}}{CH_2{=}CHCN} + H_2O$$

Polymer Production

The polymerization process parallels the emulsion process used for styrene-butadiene rubber. Either a hot or a cold process can be used, with the cold polymerization providing the same improved processing and vulcanizate properties as seen in SBR. Polymerizations are carried to 70 to 80 percent conversion and terminated to avoid gel formation. The latex must be stripped to remove unreacted butadiene and acrylonitrile.

Properties and Use

As the acrylonitrile content increases in the polymer chain, properties change predictably. The glass transition temperature increases approximately 1.5°C for each percent increase in acrylonitrile. Properties such as hysteresis loss, resilience, and low temperature flexibility will correspondingly change. The oil resistance increases with increased acrylonitrile content, as does the compatibility with polar plastics such as PVC. The major market for nitrile rubber is in the automotive area because of its solvent and oil resistance. Major end uses are for hoses, fuel lines, O-rings, gaskets, and seals. In blends with PVC and ABS, nitrile rubber acts as an impact modifier. Some nitrile rubber is sold in latex form for the production of grease-resistant tapes, gasketing material, and abrasive papers. Latex also is used to produce solvent-resistant gloves.[65] A hydrogenated nitrile rubber has been developed and commercialized that provides both better oil resistance and high-temperature resistance.[66]

CHLOROPRENE RUBBER

Chloroprene rubber (Neoprene-trade name of Du Pont) was one of the earliest synthetic rubbers, first commercialized in 1932. It has a wide range of useful properties but has not become a true general purpose synthetic rubber, probably because of its cost. It does possess properties superior to those of a number of general purpose polymers, such as oil, ozone, and heat resistance; but for these properties other specialized polymers excel. Polychloroprene thus is positioned between the general purpose elastomers and the specialty rubbers.

Monomer Production

Chloroprene monomer production starts with the catalytic conversion of acetylene to

monovinylacetylene, which is purified and subsequently reacted with aqueous hydrogen chloride solution containing cuprous chloride and ammonium chloride to give chloroprene.[67]

Production Process

Polychloroprene is produced by using an emulsion process. Two general types of processes are used: sulfur modified and unmodified. In the sulfur modified process, sulfur is dissolved in monomer and is incorporated into the polymer chain. Upon the addition of thiuram disulfide type materials and under alkaline conditions, some of the sulfur bonds are evidently cleaved to give the soluble polymer. In the unmodified process chain transfer agents are used. If neither the sulfur modified nor the chain transfer system is used, the resulting polymer is a gelled tough material. Typical polymerization systems consist of rosin acid soap emulsifier and persulfate catalyst. Conversions of 80 to 90 percent are obtained. Polymerizations are run at around 40°C. For the modified polymer, the thiuram disulfide is added after polymerization, and the latex is aged to allow the peptization reaction to occur. Acidification stops the peptization reaction. The latex is vacuum-stripped and coagulated by acoating a cold drum, with a thin frozen film continuously formed and scraped from the drum. The coagulated rubber is washed and dried.

The polymerization produces primarily *trans* 1,4 polychloroprene. The *trans* content can be increased somewhat by lowering the polymerization temperature.[68,69]

Properties and Uses

Polychloroprene is stable to oxidation and ozone. It also has flame resistance, and its oil resistance is better than that of general purpose rubbers. Its major disadvantage, other than cost, is relatively poor low-temperature properties. Because of the high stereoregularity, polychloroprene will strain crystallize, giving good tensile to unfilled stocks. At low temperatures the polymer can crystallize, making processing more difficult. Polymer made at lower temperatures will show higher unfilled tensile properties and more rapid crystallization due to the higher *trans* content (i.e., less disruption of the crystal structure). The major end uses are for conveyor belts, V-belts, hoses, and mechanical goods such as wire insulation, O-rings, and gaskets. It also has found use in single-ply roofing and adhesives.[70]

SILICONE ELASTOMERS

Silicone elastomers represent a rather unique group of polymers in that they consist of alternating silicon–oxygen bonds to form the polymer chain backbone. Side groups off the silicon atoms are selected to provide very specific properties that differentiate one type from another. The most common side group is the dimethyl structure. Replacement of small amounts of the methyl group with vinyl provide sites for cross-linking. Phenyl groups are used to improve low-temperature properties. Fluorosilicones are produced by replacing the methyl with trifluoro-propyl units. The addition of bulky phenyl side groups leads to an increase in the glass transition temperature. However, the disruption caused by such groups leads to the desired reduction or elimination of crystallization, which is critical for low-temperature properties. Such polymers have glass transition temperatures around −110°C, which is the range of the very best of the carbon-backbone polymers for low glass transition temperatures.

Monomer Production

The actual polymerization process involves a ring-opening reaction of dimethyl-substituted cyclic siloxanes. The preparation of the cyclic materials starts with the production of pure

silicon via the reduction of quartz with coke in an electric arc furnace. The silicon metal then is reacted with methyl chloride to give a mixture of silicones, from which dimethyldichlorosilane is removed by distillation.[71] Subsequent hydrolysis gives the cyclic dimethylsiloxane.

$$\underset{\text{Quartz}}{SiO_2} + \underset{\text{Coke}}{C} \longrightarrow \underset{\text{Silicon metal}}{Si}$$

$$Si + CH_3Cl \longrightarrow \underset{\text{Dimethyldichlorosilane}}{(CH_3)_2SiCl_2}$$

$$(CH_3)_2SiCl_2 + H_2O \longrightarrow \underset{\text{Cyclic dimethylsiloxane}}{[(CH_3)_2\text{—}SiO]_x}$$

Polymer Production

The polymerization process involves an equilibrium ring-opening reaction carried out in the bulk state, which can be catalyzed by acids or bases.

Uses

Silicone rubber offers a set of unique properties to the market, which cannot be obtained by other elastomers. The Si–O backbone provides excellent thermal stability and, with no unsaturation in the backbone, outstanding ozone and oxidative stability. The very low glass transition temperature, combined with the absence of low-temperature crystallization, puts silicones among the materials of choice for low-temperature performance. The fluoro-substituted versions provide solvent, fuel, and oil resistance along with the above-mentioned stability advantages inherent with the silicone backbone.

The gum polymer has rather poor tensile properties when cured, but these properties can be greatly improved by the use of silica reinforcing agents. These systems exhibit some of the greatest improvements in properties by filler addition; and because this improvement is significantly higher for silica than for other reinforcing agents, it is assumed that direct bonding occurs between the silica and the polymer. The silicone materials also may have very low surface energy, which accounts for their nonstick characteristics. Because of the inherent inertness of the materials, they have been widely used for medical purposes within the body. The largest use of silicone is in sealant and adhesive applications.

Several different methods have been developed to cure silicones. Free radical cures are possible for those polymers containing vinyl groups. The largest-volume process, however, involves room-temperature vulcanizations, which can employ either a one-component or a two-component system. In the one-component system, a cross-linking agent such as methyltriacetoxysilane is used. With exposure to moisture, hydrolysis of the cross-linking agent leads to the silanol reactive cure site; so such materials must be compounded and stored free of moisture.[72]

POLYURETHANE RUBBER

A wide range of materials is included in this class. The common feature is the use of chain extension reactions to provide products with acceptable commercial properties. The chain extension reaction effectively reduces the actual number of chain ends, thereby eliminating the generally poor properties observed when very low molecular weight polymers are cross-linked. The chain extension step involves the reaction of a difunctional polymeric polyol with difunctional organic isocyanates to give the polyurethane:

$$\text{–NCO} + \text{HO—R'—OH} \rightarrow \text{–NHCOO—R'—OCONH–}$$

The most used polyols have a polyester or a polyether backbone. A wide variety of isocyanates are used, with toluene di-isocyanate, *m*-phenylene di-isocyanate, and hexamethylene di-isocyanate the most common.

Raw Materials

The largest-volume polyether used is obtained from propylene oxide polymerized under

basic conditions:

Propyleneoxide Polypropyleneoxide

Polyester polyols are produced from a number of different materials involving diacids and diols to give the ester linkage. Aliphatic polyesters generally are used for elastomers to impart chain flexibility.

$$HO(CH_2)_2OH + HOOC(CH_2)_4COOH \longrightarrow$$

Ethylene glycol Adipic acid

Polyesterglycol

The production of isocyanates is based on the reaction of phosgene with primary amines. Toluene di-isocyanate is the most frequently used di-isocyanate.

$$CH_3C_6H_4(NH_2)_2 + COCl_2 \rightarrow$$

Toluene diamine Phosgene

$$CH_3C_6H_5(NCO)_2$$

Toluene di-isocyanate

Use

The urethane elastomers are complex segmented or block polymers. Soft, noncrystalline blocks are provided by the polyether or aliphatic polyester long chains, whereas stiff, hard blocks are produced by the reaction of aromatic di-isocyanates with low molecular weight materials such as diols or diamines. The hard blocks can phase-separate to provide a physical rather than a chemical cross-link, similar to those discussed for the butadiene styrene thermoplastic elastomers although the domain size is much smaller for the polyurethanes. Chemical cross-links can be introduced by using trifunctional materials or by adjusting the stoichiometry to allow additional reactions with the urethane or urea structures to give allophanate or biuret linkages.

Polyurethane rubbers can have high tensile strengths, excellent tear strength, and good abrasion and chemical resistance. The greatest disadvantage is the hydrolytic instability of the urethane linkage. A major use is in automotive bumpers and facia. These materials are made in a reaction injection molding process. Castable urethanes are used to produce solid rubber wheels and printing rolls. Millable urethanes can be processed on conventional rubber equipment. Unsaturation can be introduced to allow the use of conventional rubber cure systems. Cures also are possible by reaction of the active hydrogens with materials such as high molecular weight polyfunctional isocyanates that are nonvolatile at cure temperatures.

Thermoplastic polyurethane elastomers are processed by injection molding and other processes used for thermoplastics. Small gears, seals, and even automotive fender extensions can be produced by this means.

By selecting from the large number of possible reactions and stoichiometry, properties can be tailored to meet a very wide range of applications.[73-75]

PLASTICIZED POLYVINYL CHLORIDE

Generally one thinks of polyvinyl chloride as a rigid plastic—which it is, with a glass transition temperature around 85°C. The addition of polar chemicals such as dioctyl phthalate can reduce the glass transition temperature below room temperature, producing a rubbery material.

Monomer Production

Vinyl chloride is produced primarily from ethylene, which is converted to ethylene dichloride by either chlorination or oxychlorination. The ethylene dichloride is cracked to give vinyl chloride and hydrogen chloride.

$$H_2C{=}CH_2 + Cl_2 \longrightarrow ClCH_2CH_2Cl$$

Ethylene Ethylene dichloride

$$ClCH_2CH_2Cl \longrightarrow H_2C = CHCl + HCl$$

Vinyl chloride

Production Process

Polyvinyl chloride is produced by the free radical polymerization of vinyl chloride. Bulk, emulsion, solution, and suspension polymerization processes have been used.

The plasticized product can be produced by mixing the polymer and plasticizers at elevated temperatures, by dry blending in which the plasticizer is absorbed into the resin and then heated to flux, by solution blending, and by the plastisol process, in which fine polyvinylchloride powder is dispersed in the plasticizer and is relatively stable until it is heated.

Properties and Use

Plasticized polyvinyl chloride can be regarded as the first thermoplastic elastomer, as it it used in an uncross-linked form. Because of the lack of cross-linking, this material exhibits high rates of creep and stress relaxation. As with other thermoplastic elastomers, these disadvantages worsen as the temperature is increased. Although the polymer is saturated, it must be stabilized for use to prevent dehydrochlorination at processing temperatures. Because of the high chlorine content, polyvinyl chloride has excellent flame resistance as well as good electrical properties. Even at high plasticizer levels, these materials have marginal elasticity when compared to most other elastomers. Despite their shortcomings, they are used in many areas where they compete with other rubbers.

One of the larger uses is for wire and cable insulation. The flame resistance makes this the material of choice for residential wiring, extension cords, and so on. Inexpensive garden hose represents another large-volume use. Sports balls can be produced by rotational molding of plastisols. The low fabrication cost allows such products to dominate the lower-price-range market.[76,77]

MODIFIED POLYETHYLENE RUBBERS

Elastomeric polymers can be produced by the chlorination or chlorosulfonation of polyethylene. Both products start with polyethylene, either in solution or in aqueous suspension, which then is reacted to give the specified degree of substitution to obtain the desired properties. Sufficient substitution is achieved to disrupt the regularity of the polymer chain, changing it from the crystalline polyethylene plastic into amorphous elastomers.

Chlorinated Polyethylene

The chlorinated products contain around 40 percent chlorine. These materials must be stabilized with metal salts, like other chlorinated elastomers and plastics. Peroxide cross-linking generally is used. Being saturated, the materials have excellent weather and ozone resistance and can be used over a temperature range of −65°F to 300°F. The high chlorine content imparts oil resistance and relatively slow rates of burning. Typical applications where this combination of properties is required include hoses for chemical or oil resistance, tubing, and belting. In comparison with plasticized PVC, these materials have better low-temperature properties and do not suffer a loss of plasticizer because none is required.

Chlorosulfonated Polyethylene

Reaction with sulfur dioxide in addition to chlorine introduces into the polymer sites that are suitable for cross-linking. Sulfur contents in the range of 1.0 to 1.5 percent are used, with chlorine contents of 25 to 40 percent. Curing is accomplished by using metallic oxides, sulfur-bearing organic compounds, and epoxy resins. These materials have outstanding ozone resistance and show little color change upon light exposure. Good resistance to oils, heat, oxidation, weather, and corrosive materials also is exhibited. Applications include pond and pit liners, coated fabrics, light-colored roofing membranes, wire and cable insulation, chemical hose, and belting.[78,79]

REFERENCES

1. Whitley, G. S., Davis, C. C., and Dunbrook, R. F. (Eds.), *Synthetic Rubber*, John Wiley & Sons, New York, 1954.
2. Blackely, D. C., *Synthetic Rubber: Their Chemistry and Technology*, Chapter 2, Applied Science Publishers, London, 1983.
3. *Worldwide Rubber Statistics 1990*, International Inst. of Synthetic Rubber Producers, Inc., Houston, p. 6.
4. Fox, T. G., Grateh, S., and Loshaek, S., *Rheology—Theory and Applications*, Vol. 1, Academic Press, New York, 1956.
5. Billmeyer, F. W., Jr., *Textbook of Polymer Science*, John Wiley & Sons, New York, 1984.
6. Odian, G., *Principles of Polymerization*, pp. 19–24, McGraw-Hill, New York, 1970.
7. Flory, P. J., *Principles of Polymer Chemistry*, Chapter 7, Cornell University Press, Ithaca, NY, 1953.
8. Rodriguez, F., *Principles of Polymer Systems*, pp. 138–142, McGraw-Hill, New York, 1982.
9. Rudin, A., *The Elements of Polymer Science and Engineering*, p. 126, Academic Press, New York, 1982.
10. Ibid., p. 402.
11. Krause, S., *J. Macromol. Sci.*, **C7** (2), 251 (1972).
12. Morton, M., and Fetters, L. J., in *Polymerization Processes*, C. Schildknecht and I. Skeist, (Eds.), Chapter 9, John Wiley & Sons, New York, 1977.
13. Brandrup, J., and Immergut, E., *Polymer Handbook*, Vol. II, pp. 105–387, John Wiley & Sons, New York, 1975.
14. Ref. 7, pp. 178–195.
15. Biesenberger, J. A., and Sebastian, D. H., *Principles of Polymerization Engineering*, Chapters 1 and 2, John Wiley & Sons, New York, 1983.
16. Ref. 7, Chapter 11.
17. Oberster, A. E., Bouton, T. C., and Valaitis, J. K., *Die Angewandte Markomoleculare Chemie*, **29/30**, 291 (1973).
18. Florey, P. J., *J. Chem. Phys.*, **17**, 223 (1949).
19. Natta, G., in *Polymer Chemistry of Synthetic Elastomers*, Part I, J. Kennedy and E. Tornqvist (Eds.), Chapter 7, John Wiley & Sons, New York, 1969.
20. Evans, C., *Practical Rubber Compounding and Processing*, Applied Science Rubbers, Englewood, NJ, 1981.
21. Semegen, S. T., and Cheong, S. F., *Vanderbilt Rubber Handbook*, pp. 18–41, R. T. Vanderbilt Company, Inc., 1978.
22. Morton, M., *Rubber Technology*, Van Nostrand Reinhold Co., New York, 1973.
23. Webster, C. C., and Baulkwill, W. J., Chapters 1 and 11, *Rubber*, John Wiley & Sons, New York, 1989.
24. Ref. 13, Vol. III, p. 54.
25. Treloar, L., *The Physics of Rubber Elasticity*, Clarendon Press, Oxford, 1949.
26. Poh, W. N., *Elastomers*, p. 12 (1989).
27. *Synthetic Rubber: The Story of an Industry*, International Inst. of Synthetic Rubber Producers, New York, 1973.
28. Stavely, F. W., *Ind. Eng. Chem.*, **48**, 778 (1956), presented at Div. Rubber Chem., ACS, Philadelphia, PA, 1955.
29. Eng. Patent 827365, to Goodrich-Gulf Chem., Dec. 1954, priority data.
30. Ref. 3, pp. 62–63.
31. Ref. 3, p. 77.
32. Ref. 3, p. 19.
33. Renninger, T. J., presentation to International Rubber Study Group, Ottawa, Sept. 1990.
34. Faith, W. L., et al., *Industrial Chemicals*, p. 731, John Wiley & Sons, New York, 1957.
35. Ref. 21, p. 53.
36. Dunbrook, R. F., *India Rubber World*, **117**, 203–207 (1947).
37. Gardon, J. L., *Rubber Chemistry and Technology*, **43**, 74–94 (1970).
38. Storey, E. B., *Rubber Chemistry and Technology*, **34**, 1402 (1961).
39. Kuntz, I., *J. Poly. Sci.*, **54**, 569 (1969).
40. Bouton, T., and Futamura, S., *Rubber Age*, **3**, p. 33 (1974).
41. Mostert, S., and Van Amergongen, G., British Patent 1,136,189, 1968.
42. Ref. 19, Chapter 6.
43. Day, G., and Moore, D., paper at 26th Annual Meeting, IISRP, May 1985.
44. Oshima, N., Salcacobore, M., and Tsutsumi, F., paper at 27th Annual Meeting, IISRP, May 1986.
45. Ref. 3, p. 9.
46. Ponder, T., *Hydrocarbon Processing*, **55** (10), 119–121 (1976).
47. Womeldorph, D., *Am. Petrol. Inst.* (May 14, 1958).
48. Welch, M., *Hydrocarbon Processing*, **57** (11), 131–136 (1978).
49. Adams, H., Farhat, K., and Johnson, B., *Ind. Eng. Chem., Prod. Res. Dev.*, **5** (2), 127 (1966).
50. Halasa, A., Schulz, D., Tate, D., and Mochel, V., in *Advances in Organometallic Chemistry*, F. Stone and R. West (Eds.), Vol. 18, Academic Press, New York, 1980.
51. Tate, D., and Bethea, T., *Encyclopedia of Polymer Science and Engineering*, Vol. 2, pp. 568–72, John Wiley & Sons, New York, 1985.

52. Lauretti, F., and Gargani, L., 27th Annual Meeting, IISRP, June 1987.
53. DeChirico, A., Lamzani, P., Eaggi, E., and Bruzzone, M., *Makromol. Che.*, **175**, 2029 (1974).
54. Buckenell, C., "Toughened Plastics," *Applied Science*, London (1977).
55. Natta, G., *J. Poly. Sci.*, **51**, 411 (1961).
56. Lukach, C., and Spurlin, H., *Copolymerization*, G. Ham (Ed.), Interscience, New York, 1964.
57. Crespi, G., and DiDrusco, G., *Hydrocarbon Processing*, **48**, 103–107 (Feb. 1969).
58. Vandenberg, E., and Repka, B., in Ref. 12, Chapter 11.
59. Scagliotti, F., Milani, F., and Galli, P., paper at 26th Annual Meeting, IISRP, May 1985.
60. Miles, D., and Briston, J., *Polymer Technology*, p. 299, Chemical Publishing Company, New York, 1963.
61. Kennedy, J., in *Polymer Chemistry of Synthetic Elastomers*, Part I, J. Kennedy and I. Tornqvist (Eds.), Chapter 5, John Wiley & Sons, New York, 1969.
62. "Butyl Rubber Reaches 50 Year Mark," *Elastomerics*, pp. 30–31 (Mar. 1988).
63. Konrad, E., and Tschunkur, E., U.S. Patent 1,973,000 to I. G. Farbenindustrie, 1934.
64. Ref. 3, p. 9.
65. McKenzie, G., paper at 27th Annual Meeting, IISRP, May 1986.
66. Hashimoto, K., et al., paper at 26th Annual Meeting, IISRP, May 1985.
67. Ref. 1, p. 770.
68. Hofmann, W., Ref. 60, Chapter 4.
69. Ref. 2, pp. 175–194.
70. Brodrecht, L., *Chemical Economics Handbook*, SRI International, 1989.
71. Polmateer, K., *Rubber Chemistry and Technology*, **16** (3), 470 (1988).
72. Noble, M., Ref. 21, pp. 216–232.
73. Saunders, J., Ref. 19, Chapter 8.
74. Bakus, J., Ref. 12, Chapter 17.
75. *Handbook of Elastomers*, pp. 643–659, Marcel Dekker, Inc., New York, 1988.
76. Penn, W., *PVC Technology*, Applied Science Publishers, London, 1971.
77. Ref. 2, Chapter 8.
78. Broadrecht, L., Mulach, R., and Tauchiya, K., in *Chemical Economics Handbook, Elastomers*, SRI International, Menlo Park, CA, 1989.
79. Johnson, J., Ref. 21, pp. 295–307.

19

Synthetic Resins and Plastics

Jerome K. Knapczyk* and Robert H. M. Simon**

INTRODUCTION

The emergence of the synthetic resins and plastics industry from primitive beginnings a century ago to an international multibillion dollar per year industry today has profoundly changed the way people live. Plastics can perform in ways that are not possible with traditional materials and therefore have become an essential component in our technologically based society. For many applications, such as videotape and heart valves, plastics are the only feasible material. In the United States alone, 61.5 billion lb of resins were sold in 1990.[1] The upward trend in U.S. production of resins is illustrated in Fig. 19.1.[2] The expectation of continued consumer demand, as well as industry profits, supports a large research and development effort that promises even better materials for the future.

Though most plastic and resin products are derived from petroleum, they account for less than 3 percent by weight of total consumption of oil and gas for both material and energy consumption.[3] The amount of energy required to fabricate a finished plastic is generally less than that required for materials such as aluminum, steel, and glass. Weight and design advantages result in further energy savings. Formability means fewer parts, fewer fasteners, and fewer steps needed to put together complex assemblies such as auto dashboards, airplane seats, and boat hulls. This translates directly into lower production cost. Moreover, a lighter product means less fuel consumption when the finished product is transported or finds its way into cars or airplanes, for example.

The commercial history of synthetic resins and plastics can be traced to about 1869, when John Wesley Hyatt and his brother Isaiah, who were seeking a substitute for ivory, developed a practical process for converting cellulose nitrate into useful products.[4] It was mixed with camphor and molded into dentures, billiard balls, toothbrushes, combs, dolls, and collars. This material, which was called celluloid, was one of the developments that made the early motion picture industry possible, as it also could be cast into transparent films of good optical quality. Because of its flammability and poor dur-

*Monsanto Chemical Company, Springfield, MA 01151.
**Monsanto Chemical Company (retired).

Fig. 19.1. U.S. resin sales, 1940–90.

ability, celluloid has been largely replaced by other plastics, though it is still used for table tennis balls.

During the first decade of the twentieth century, Leo Baekeland developed the first practical process for making molded objects from phenol-formaldehyde resins. The company he formed was named General Bakelite Co. and he called his product bakelite. The development of these early materials was largely an art because very little was known about the chemistry and physical changes that led to the final products. In the mid-1920s Hermann Staudinger hypothesized that plastics owe their properties to the very large size of the molecules that compose them. Then, within ten years, contributions by Wallace Carothers, Herman Mark, Paul Flory, and others led to the incontrovertible conclusion that the chemistry and physical dynamics that lead to the formation and characteristic properties of polymers are the same as those for smaller molecules when account is made for size and complexity.[5] This very fundamental understanding has led to the synthesis and modification of thousands of new polymeric materials and to the development of the modern plastics industry.[6]

The word "plastic" was originally used as an adjective to denote formability. In the 1920s it began to be used as a noun to refer to synthetic materials such as bakelite and celluloid. Now it is used generically to refer to a large class of manufactured organic polymeric materials (Fig. 19.2). The American Society for Testing Materials (D 1695–77) has defined a plastic as "a material that contains as an essential ingredient an organic substance of large molecular weight, is solid in its finished state, and, at some stage in its manufacture or in its processing into finished articles, can be shaped by flow."

Though the words "plastics" and "resins" sometimes are used interchangeably, workers in the polymer industry generally make a distinction between them. Originally "resin" referred to the natural exudates of plants and trees such as rosin, copal, amber, elemi, kaure, manila, mastic, and batu. Later the word also became to be used for resin-like substances of animal origin such as shellac and casein.

In the synthetic polymer field, "resin" also refers to unblended, synthetic polymeric raw materials from which a finished adhesive, coating, or plastic product is made. In the manufacture of a polyethylene irrigation pipe, for example, the raw product synthesized from ethylene monomer is called a resin. It is considered a plastic material after the resin is blended with fillers, pigments, antioxidants, or other substances, and then extruded.

Besides plasticizers and stabilizers, other additives used in plastics include other resins, minerals, dyes, fibers, biocides, processing

Fig. 19.2. A classification of polymeric substances.

aids, flame retardants, antistats, ultraviolet absorbers, rubbers, or any other substance needed to meet the requirements of the manufacturing process and the product's end use.[7]

The ability, or inability, of a resin to react under certain conditions is an important property. Many resins, such as polyethylene, are relatively unreactive and are used primarily for the physical properties they impart to the finished product. Other resins are synthesized with reactive sites so that they can be made to react further. Ion exchange resins fall into this category; so do the polymeric components of reactive coatings, and adhesives, and reactively processed elastomers and plastics.

Resins and plastics also can be combined with inorganic substances to make hybrid materials that combine the desirable properties of the dissimilar materials. Examples include safety glass (glass/polyvinylbutyral/glass laminate), carbon fly rods (carbon fiber/epoxy resin composite), snack food packaging (aluminum/multilayer plastic film laminate), and computer circuit boards (glass cloth/polyimide composite).[8]

Plastics and resins can be classified in several different ways. They can be grouped according to chemical composition (polyolefin, polyacrylate); according to physical characteristics (elastomeric, thermoplastic, thermoset); according to form (sheet, film, fiber); and according to function (packaging film, structural adhesive, engineering thermoplastic).

Although these classifications are useful, it often is difficult to make a clear distinction between chemistry, form, and function. The distinctions often are arbitrary, and consequently the literature often is confusing. For example, "plastic" sometimes is used generically to cover all solid polymer products whether they are rigid or not. However, the term "elastomer" often is reserved for polymeric materials with elastic properties. Elastomers sometimes are defined as polymers that can be stretched to at least twice their

original length, and, when the stress is withdrawn, will rapidly retract to the original length.[9] When a distinction is made between plastics and elastomers, "plastics" refers to less resilient organic materials used for their structural rigidity.

This chapter will discuss some fundamental characteristics of resins and plastics technology. It will cover some important aspects of structure, properties, and processing. More thorough discussions of elastomeric materials, adhesives, coatings, and textiles will be found in other chapters.

POLYMER STRUCTURE AND NOMENCLATURE

Plastics and resins are polymeric substances derived from monomeric precursors. The monomeric units can be put together in many ways. The chain can be linear (even though it may be tightly coiled), branched, or cross-linked (Fig. 19.3). When the polymer is derived from a single monomer, it is called a homopolymer. When the polymer contains two or more different monomeric components, it is called a copolymer. Depending on the monomers and other reactants used in the synthesis, and the polymerization conditions, the monomeric units can be arranged in various ways. They can alternate randomly or in a regular way. They also can be arranged in blocks and in branches that radiate from the main chain (Fig. 19.4).

A polymer is named by enclosing within parenthesis the name of the repeating unit, or name of the monomer(s), from which it is derived, preceded by the prefix poly. The

Fig. 19.3. Some polymer chain configurations.

```
 -A-A-A-A-A-A-A-A-      Linear Homopolymer

 -A-B-A-B-A-B-A-B-      Linear Alternating Copolymer

-A-A-A-A-A-B-B-B-B-B-   Linear Diblock Copolymer

-A-A-A-A-A-A-A-A-A-A-   Graft Copolymer
        |
        B
        |
        B
        |
        B
        |
        B

-A-A-A-A-A-A-A-A-A-A-   Linear Side-Chain Homopolymer
   |       |     |
   A       A     A
   |       |     |
   A       A     A

             A
             |
             A
             |
-A-A-A-A-A-A-A-A-A-A    Branched Homopolymer
   |
   A
   |
   A
   |
 A-A-A-A-A-A-A-A-
             |
             A
             |
             A
             |
             A
```

Fig. 19.4. Some polymer backbone structures.

systematic name, or name approved by the International Union of Pure and Applied Chemistry (IUPAC), is derived from the IUPAC designation of the repeating units.[10] Trivial, semisystematic, or trade names often are used, and for simple polymers the parenthesis frequently is dropped. Table 19.1 gives names and a structural representation for a selection of polymers.[11]

Sometimes polymers with identical repeating units have different common names. This reflects different modes of synthesis. For example, both polyethylene and polymethylene have the same backbone structure, which can be represented by either —(CH_2CH_2)— or —(CH_2)—. By convention, however, the former is used to signify a polymer prepared from ethylene, whereas the latter is used for a polymer made from a monomer with only one carbon, such as diazomethane. Polymers like these, with the same repeating units, but made in different ways, can have quite different physical properties. The name given to a polymer says something about its repeating units, or the molecules from which it is synthesized, but it does not characterize it with respect to molecular weight, polydispersity, stereochemistry, or chemical and physical properties.

Unsymmetrical monomers such as propylene link together in what is called a head-to-tail manner; that is, regions such as

$$-CH_2\underset{|}{C}HCH_2\underset{|}{C}HCH_2\underset{|}{C}HCH_2\underset{|}{C}H-$$

will predominate over regions such as

$$-CH_2\underset{|}{C}H\underset{|}{C}HCH_2CH_2\underset{|}{C}H\underset{|}{C}HCH_2-$$

Unsymmetrical monomers can produce polymers with various stereochemical arrangements. When the carbons bearing the methyl group in polypropylene, for example, have the same configuration, the polymer is called isotactic; when alternate carbons with the methyl group have the same configuration, the polymer is called syndiotactic; and when there is a random distribution of configurations, the polymer is called atactic (Fig. 19.5). The stereochemistry of a polymer will strongly influence its properties. When a thermoplastic polymer is stereoregular, the individual polymer strands can more easily align in an orderly way. Thus a stereoregular polymer will tend to be more crystalline than an atactic one.

Molecular weights can be very low or very high, depending on the degree of polymerization. When the degree of polymerization is relatively small, usually less than 100, the polymer is called an oligomer. In practice it is very difficult to carry out a polymerization so that all the molecules produced have exactly the same molecular weight. Thus, there is a distribution of molecular weights in a polymer.

Unlike nonpolymeric molecules, the molecular weight obtained for a polymer will be an average one, and will depend on the way that molecular weight is measured. Molecular weight determinations that depend upon counting numbers of macromolecules of each weight, such as by chromatographic, osmotic pressure, boiling point elevation, and melting point depression methods, produce what is called number average molecular weights, M_n. If light scattering or centrifugation methods are chosen to obtain molecular weights, larger

TABLE 19.1 Nomenclature of Some Common Monomers and Polymers

Monomer	*Repeating Unit*	*Names*
Ethylene	$(-CH_2CH_2-)$	Polyethylene
Propylene	$(-CH(CH_3)CH_2-)$	Polypropylene Poly(propylene)
Styrene	$(-CH(C_6H_5)CH_2-)$	Polystyrene Poly(1-phenylethylene)
Vinyl chloride	$(-CH(Cl)CH_2-)$	Polyvinyl chloride Poly(vinyl chloride) Poly(1-chloroethylene)
Vinylidene chloride	$(-CCl_2CH_2-)$	Polyvinylidene chloride Poly(vinylidene chloride) Poly(1,1-dichloroethylene)
Acrylonitrile	$(-CH(CN)CH_2-)$	Polyacrylonitrile Poly(1-cyanoethylene)
Methyl methacrylate	$(-C(CH_3)(COOCH_3)CH_2-)$	Poly(methyl methacrylate) Poly[1-(methoxycarbonyl)-1-methylethylene]
Formaldehyde	$(-CH_2O-)$	Polyformaldehyde Poly(oxymethylene)
Caprolactam	$[-NHC(=O)(CH_2)_5-]$	Nylon 6 Polycaprolactam Poly(ε-caprolactam) Poly[imino(1-oxohexamethylene)]

Polypropylene configurations.

Fig. 19.5. Homopolymer backbone configurations resulting from addition polymerization of an unsymmetrical monomer.

size molecules make a bigger contribution to the average value obtained. The molecular weight obtained by these techniques is called weight average molecular weight, M_w. Weight average molecular weights are always larger than number average molecular weights. In the rare cases where all of the molecules have the same molecular weight, the ratio M_w/M_n is unity. This ratio, called the polydispersity index, is a measure of the molecular weight distribution in a polymer. The greater the range of molecular weights in a polymer, the larger will be the index.

Molecular weight and polydispersity often can be manipulated. The degree of control over these properties will depend on the polymerization mechanism, conditions during polymerization, and final processing conditions. Polymers with low polydispersity have higher and sharper melting points than those

with large polydispersity. Other properties such as clarity, hardness, melt flow behavior, solvent resistance, and strength also depend on polydispersity.

PROPERTIES OF RESINS AND PLASTICS

There are unlimited combinations of polymer structure, molecular weight, and conformation possible, and each has a unique set of chemical and physical properties. For this reason properties of a given polymer type can vary greatly. Most plastics are used for their material properties and not for the chemical reactions that they might undergo. Important material properties include density, melting point, glass transition temperature, rheology, hardness, abrasion resistance, conductivity, tensile strength, gas permeability, chemical and radiation resistance, color, clarity, surface energy, and so on.

Polymers would be low-viscosity fluids if they were not highly tangled or otherwise bonded together. When intermolecular interactions are weak (e.g., entanglement, hydrogen bonds, dipole–dipole interactions, and Van der Waals bonds), the polymer is called a thermoplastic. When the entire polymer matrix is linked together by covalent bonds into a single three-dimensional network, the polymer is called a thermoset. Thermosets are said to be cross-linked.

If a thermoplastic solid is heated to sufficiently high temperatures, the bonds that make it rigid at lower temperatures begin to weaken. Individual chains start to untangle and slip by each other, and the mass begins to soften and flow. The plastic can be cast, extruded, or thermoformed at this stage. When the viscous shaped material cools, the mass becomes rigid again, as intermolecular bonds re-form. In principle this cycle can be repeated indefinitely, but in practice thermal and oxidative degradation may take place at the elevated temperatures needed for processing. This limits the number of cycles that can occur before the plastic's desirable properties are altered.

At sufficiently high temperatures all organic materials degrade as bonds begin to fragment, rearrange, or oxidize. Even at ambient conditions, oxidation and other changes caused by photochemical reactions, or by moisture, acids, and other contaminants in the air, may occur. These aging processes often are accompanied by cross-linking, the formation of lower molecular weight substances, and colored by-products. Over a period of time the product may weaken, become brittle, and change color. Thus, many plastics are formulated with antioxidants and light stabilizers to minimize degradation during processing and to promote durability during use.

In most cases, the weak intermolecular bonds of thermoplastics allow them to be dissolved in a solvent or blended with other substances. A solution of thermoplastic in a volatile solvent is called a lacquer. Higher molecular weight nonvolatile, nonreactive components such as plasticizers or other resins also may be soluble in a thermoplastic to some degree. Plasticizers dilute a resin and reduce intermolecular forces between molecules. They are added to soften a resin or make it flexible, or to make it easier to process. Vinyl upholstery, some films used in packaging, and many other plastics are made with high levels of plasticizer.

The weak intermolecular forces that make thermoplastics easy to process can lower resistance to abrasion, chemical attack, or etching by a solvent, and may restrict their use at higher temperatures where they can soften or creep. In contrast, thermoset plastics do not melt or dissolve in a solvent, and they often are more resistant to thermal deformation and chemical attack than are thermoplastics. The degree of rigidity under thermal and mechanical stress is related to the amount of cross-linking. Highly cross-linked thermosets retain their shape at elevated temperatures and under strong deformation forces, whereas lightly cross-linked thermosets are easily deformed or elastomeric.

Molecular order in a thermoplastic depends on intrinsic factors such as stereoregularity, branching, polydispersity, and polarity of the repeating units, and, to some degree, on extrinsic factors such as the polymer's purity and the way it is isolated and processed.

Highly ordered arrangements are said to be crystalline. Crystallinity in a polymer is unlike that for low molecular weight organic compounds such as sucrose; complete crystallinity seldom is attained. Instead, crystalline order occurs in domains called crystallites, which are regions of the polymer matrix where orderly packing of the individual polymer strands takes place. Crystallites alternate with highly disordered, or amorphous, regions (Fig. 19.6). Such polymers are referred to as semicrystalline. Because the size of the crystalline regions may be large enough to scatter light, and their refractive index is generally different from that of the amorphous regions, semicrystalline polymers such as polyethylene often are translucent or opaque.

The distribution of the amorphous and crystalline regions in a polymer can be random (Fig. 19.6a) or oriented (Fig. 19.6b). When semicrystalline thermoplastics are drawn into filament, or stretched into film at elevated temperatures, the long axes of the crystallites align in the direction of drawing (Fig. 19.6b). If this orientation is frozen in by quenching the hot drawn polymer, the orientation is set, and some material properties become anisotropic. Properties such as refractive index, elongation, and tensile strength will differ in the polymer's three dimensions. In practice, the tensile strength in the draw direction is improved by one or two orders of magnitude; and this improvement is used to advantage in the production of packaging film and fiber, where enhanced strength is needed to compensate for the small cross-sectional areas of the product.

In general, highly crystalline thermoplastics have a higher density, are more resistant to solvents, and melt more sharply at higher temperatures than amorphous polymers or polymers with low crystallinity. Although highly crystalline polymers have greater stiffness, they usually are more brittle or less resilient to shock than amorphous polymers. Thus, in applications where stiffness and solvent resistance are important, a highly crystalline polymer is a likely choice, but where clarity and resilience are needed, an amorphous polymer might be used. Often a balance between amorphous and crystalline regions is needed to produce the combination of strength and flexibility required for a given application.

Fig. 19.6. Crystalline and amorphous regions in a semicrystalline thermoplastic polymer. (a) Unoriented. (b) Oriented.

Another property that determines a plastic's usefulness is its modulus. Modulus is related to stiffness; high stiffness corresponds to high modulus. Modulus is defined as the ratio of stress (deforming force per unit cross-sectional area) to strain (increment of deformation) for elastic deformation (Fig. 19.7). Its dimensions are the same as those of pressure or tensile strength.

The modulus obtained for a material depends on the way that it is measured. When the term "modulus" is used without qualifiers, the slope of the steepest part of the stress–strain

Fig. 19.7. Stress–strain relationship. The initial slope defines the modulus.

Fig. 19.8. Volume–temperature changes for semicrystalline (*A*) and amorphous (*B*) polymers.

curve, or initial modulus, is meant. Tensile modulus corresponds to the initial slope of the stress–strain curve generated when a specimen is clamped at one end and pulled at the other. Torsional modulus is the torque required to twist a specimen clamped at one end through a specified angle. Modulus usually decreases with plasticizer content and increases when fillers and reinforcing fibers are added.

A polymer's modulus may strongly affect its glass transition temperature, T_g, which is the temperature at which segments of the polymer chain, about 10 to 30 units long each, acquire enough energy to move around, about fixed positions in the matrix. Below T_g a polymer is rigid or glasslike. Above T_g, segmental motion sweeps out increasingly larger volumes than the volume of the atoms that make up the segments (Fig. 19.8). As this so-called free volume increases, the polymer's density and modulus decrease.[12] In addition to its becoming less rigid, other properties such as impact strength and elongation also change substantially.

In contrast to small crystalline molecules, where changes from a solid to a liquid take place over a narrow temperature span, amorphous polymers do not exhibit well-defined melting points. As temperature increases above T_g, the polymer becomes less rigid as the chains become more and more mobile and begin to slip by each other. Viscosity decreases gradually, and the polymer eventually begins to flow.

A semicrystalline polymer is more rigid above T_g than a comparable amorphous polymer. For example, a crystalline polymer such as polyethylene, with a T_g well below ambient temperature, will retain useful structural properties above ambient temperature. Crystallites serve as inter-chain crosslinking sites that retard the mobility of polymer chains above T_g until the crystallites begin to melt, at a temperature called the crystalline melting point, T_m. Table 19.2 lists T_g of some common polymers.

A polymer's T_g and T_m can be determined in several ways. A convenient method is thermal analysis. Measurements are carried out in a calorimeter that scans temperature and measures the quantity or rate of heat absorbed by a substance compared to a standard. Scans for an amorphous and a semicrystalline polymer are shown schematically in Fig. 19.9.

The influence of T_g on the mechanical properties of polymers can be dramatic. Properties such as refractive index, abrasion and solvent resistance, thermal conductivity, and response to mechanical shock may show step changes at T_g. On heating through T_g, the modulus of an amorphous polymer may decrease by two to five orders of magnitude.

In Fig. 19.10, the effect of increasing the molecular weight of a hypothetical amorphous oligomer is to cause a shoulder in the modulus

TABLE 19.2 Approximate Glass Transition Temperatures (T_g) for Some Common Polymers

Polymer	T_g(°C)
Poly(dimethyl siloxane)	−123
Polyethylene (LDPE)	−122
Polybutadiene	−85
cis-Polyisoprene	−70
Poly(butyl acrylate)	−63
Poly(propylene oxide)	−62
trans-Polyisoprene	−53
Polychloroprene	−50
Poly(ethylene adipate)	−48
Polypropylene (atactic)	−20
Poly(vinyl acetate)	28
Poly(ethylene terephthalate)	70
Poly(vinyl chloride)	80
Polypropylene (isotactic)	100
Polystyrene	100
Poly(methyl methacrylate)	105
Polyacrylonitrile	105
Poly(tetrafluoroethylene)	115
Cellulose triacetate	197

curve at about 10^6 dynes/cm^2. As molecular weight increases, a plateau of increasing height and length is produced, as entanglement begins to retard slippage between polymer chains. The temperature–modulus plateau at $T > T_g$ corresponds to a state where a polymer exhibits the resilient properties of a viscous solid, often referred to as viscoelastic

Fig. 19.9. Differential thermal scan showing discontinuities in heat absorption ($\Delta H/g$) for the semicrystalline (A) and amorphous (B) polymers in Fig. 19.8.

Fig. 19.10. Generalized modulus–temperatures curves for low (A), moderate (B), and high (C) molecular weight amorphous thermoplastic polymers.

properties.[13] These properties are discussed further below, in the section on "Polymer Rheology."

The effect of temperature on the modulus of a semicrystalline resin is shown in Fig. 19.11. The modulus change is less dramatic near T_g because backbone rigidity is provided by the crystallites. When the temperature becomes high enough to melt the crystallites, the polymer is free to flow.

Figure 19.12 compares the modulus–temperature response of an amorphous thermoplastic polymer with thermosets that are lightly and highly cross-linked. Cross-links in thermosets do not melt out like crystallites in semicrystalline thermoplastics; so the modulus–temperature curve is flatter up to the point of thermal decomposition. Modulus above T_g is strongly dependent on the number of cross-links per unit volume, or

Fig. 19.11. Modulus–temperature curve for a semicrystalline polymer.

Fig. 19.12. Generalized modulus–temperature curve for an amorphous thermoplastic (*A*) and for thermosets that are lightly (*B*) and highly (*C*) cross-linked.

molecular weight between cross-links. For example, if the molecular weight between cross-links is between 2500 and 5000, modulus will be in the vicinity of 5 to 10×10^6 dynes/cm^2. This is the range for elastomers where there are large threadlike segments to coil and stretch. The glass transition temperatures of flexible thermosets such as rubbers, sealants, and pressure-sensitive adhesives are below room temperature. In phenol-formaldehyde plastics and some other thermosets, where the degree of cross-linking is very high, a modulus greater than 10^8 dynes/cm^2 may be exhibited above T_g.

T_g depends on some of the same chemical and steric factors that affect crystallinity. It also depends to some degree on how it is measured. When obtained with oscillating stress–strain response methods, T_g will increase with the rate of applied stress (Fig. 19.13).

As might be expected, other viscoelastic properties will show a time–temperature dependency. For example, if an indenter is used to measure the hardness of a thermoplastic, the time that the indenter is impressed on the surface is a measurement variable. Over a long impression time the polymer may yield under the stylus point, allowing it to sink deeper into the surface. Thus the polymer will appear to be less hard than it seems if the impression time is short. This time–property relationship also is exhibited during impacts. At temperatures above T_g a polymer might exhibit ductile properties when deformed at low rates and brittle fracture at high deformation rates.

Fig. 19.13. Generalized modulus–temperature curves showing the effect that periodic strain rate has on T_g.

The creep behavior of a plastic is an important consideration in its use. Under static loads thermoplastics may deform at temperatures below a T_g measured under dynamic conditions. A useful way to express creep behavior is by what is called the heat deflection temperature (HDT). HDT is the minimum temperature that results in a specified deformation, usually 0.01 inch under a static load, usually 1.82 or 0.46 N/mm^2 (264 or 66 psi, respectively).[14] For an unreinforced plastic, this is around 20°C below the T_g determined at a frequency of one oscillation per second.

IMPORTANT CLASSES OF PLASTICS AND RESINS: THERMOPLASTICS

Because of their versatility and processing ease, thermoplastics account for about 85 percent of the domestic resins market of about 61 billion lb/year.[1] Some of the more important members of this class of resins are discussed below.

Polyolefins

Polyolefins, particularly polyethylene and polypropylene, are the largest-selling class of resins. Domestic sales volume in 1990 for polyethylene was 20 billion lb, and for polypropylene and its copolymers it was 8 billion lb.[1] The large volume of polyolefins sold reflects a broadly useful property spectrum, ease of fabrication, and low cost. Lower cost is a result of a favorable raw material supply and continuing improvement in manu-

facturing processes. (See below, "Polymerization Methods" in section on "Polymer Synthesis.")

Polyethylene. Polyethylene accounts for about 32 percent of domestic resin sales. It is a high molecular weight, semicrystalline polymer prepared from ethylene. Depending on its manufacturing process, its degree of polymerization will vary from about 500 to 500,000. The monomer is prepared in large quantities by the cracking of natural gas and light petroleum.

Polyethylene was first manufactured in England in 1939.[15] The first applications of polyethylene—which took advantage of its relative inertness, moisture resistance, and good electrical insulating properties—were in marine and radar cable insulation. At room temperature polyethylene is resistant to most chemicals except strong oxidants. Currently, the largest uses for polyethylene are for containers and packaging film. This includes kitchen food wrap and store packaging for fruits, vegetables, and other food products, and heavy-duty sacks for fertilizers and other bulk products.

Commercial polyethylene is available in various molecular weights and branching, resulting in different mechanical properties and processibility. The various grades can be divided into two principal categories, low density polyethylene (LDPE) and high density polyethylene (HDPE). The density of LDPE varies from about 0.90 to 0.94 g/cm^3, and HDPE varies from 0.95 to 0.97 g/cm^3. Density is related inversely to the average frequency and length of molecular chain branching. As polyethylene density increases, impact strength and elongation tend to drop off, while tensile strength, abrasion resistance, hardness, rigidity, and heat distortion temperature increase.

There are two types of low density polyethylene, a highly branched type made by a high pressure process (HP-LDPE) and one prepared at much lower pressures that has higher molecular weight, less branching, and fewer periodic short side chains than the former type. The latter is called linear low density polyethylene (LLDPE).

HP-LDPE is prepared at pressures of 1000 to 3000 atmospheres and at temperatures from about 150 to 250°C. Polymerization takes place by a free radical mechanism initiated by molecular oxygen or peroxides such as benzoyl peroxide. If lower pressures were used, the polymer formed this way would have an excessively low molecular weight. This is so because molecular weight is determined largely by the rates of chain growth and chain termination, and for the growing polyethylene radical the rate of termination is very high. High pressures increase the concentration of reactants and the probability that the growing chain will encounter a monomer before termination can occur.

Elevated reaction temperatures and highly reactive propagating free radicals result in competing chain transfer reactions. Chain transfer occurs when a hydrogen atom is picked off the backbone of a growing intermediate by another free radical center (see below, under "Free Radical Addition Polymerization"). Because chain transfer reduces molecular weight and causes branching, the polymer made this way is highly branched. Branching prevents close packing, making crystallization more difficult. Typical HP-LDPE exhibits crystallinity of about 20 to 55 percent, compared to 60 to 90 percent for unbranched polyethylene. Control over molecular weight to produce various grades of LDPE is carried out by adding chain transfer agents to the monomer feed.

LDPE accounts for about 55 percent of U.S. sales of polyethylene. It supplies 41 percent of the 6.5 billion lb polyethylene film market,[16] 53 percent of the 850 million lb injection molding market, and the majority of the extrusion, wire-and-cable, and shrink-wrap markets.[17] HP-LDPE is widely used for higher-clarity applications. Its low melting point allows it to be coated or extruded at relatively low temperatures and higher rates than LLDPE. Extrusion coatings on paper and board are used to provide moisture resistance. Coated milk cartons and multiwalled paper bags are examples. It is used for

power and cable insulation, and is extruded into irrigation piping and monofilament for rope, nets, and lawn chairs. It is injection-molded into buckets, bottles, drums, baskets, toys, and laboratory ware.

HDPE and LLDPE are prepared at lower pressures by non–free radical chemistry. The Phillips process, for example, uses proprietary chromium or molybdenum oxide catalysts on finely divided silica or silica–alumina supports.[18] Polymerization occurs between 70 and 200°C and at pressures in the range of 30 to 40 atmospheres.

A competing low pressure process uses what are called Ziegler-Natta catalysts.[19] Polymerization with these catalysts take place at temperatures below 100°C and pressures in the neighborhood of 2 to 4 atmospheres. These catalysts were developed in the early 1950s through the discoveries of Karl Ziegler in Germany and Giulio Natta in Italy, and their importance brought Ziegler and Natta the Nobel Prize in chemistry in 1963. Since then many related polymerization catalysts have been prepared.[20]

Both heterogeneous and soluble Ziegler-Natta catalysts are used. They are formed by combining compounds of elements from Group IV to VII with organometallics from metals of Groups I and III. Typically they are combinations of aluminum alkyls and aluminum alkyl halides with Ti(III) and Ti(IV) halides. They can be adsorbed on finely divided mineral supports to provide a large surface area for heterogeneous processes. Though some of the details of how these catalysts function are not known, it is generally accepted that monomers form pi-complexes with the transition metal component and that additional monomers are inserted, one after another, into the polarized metal–carbon bond. The polymer grows from the catalytic site much as a blade of grass grows from its root.

Transition metal catalysts produce polyethylene resins with less branching than polyethylene made by high pressure processes. In turn, the polymer molecules can pack closer together, resulting in higher density and crystallinity. Crystallinity reduces clarity and increases modulus and tensile strength. The melting point of HDPE is about 20°C higher than that of LDPE.

HDPE accounts for about 42 percent of the domestic polyethylene market.[1] It is harder, more rigid, and less waxy in appearance than HP-LDPE. Its higher density gives it better gas barrier properties. Because it softens above 100°C, containers made with HDPE can be sterilized, but those made with LDPE cannot. It is extruded into pipe, trash bags, food packaging, multiwalled sack liners, wire and cable coverings, and thick industrial film. A large application is for blow-molded containers and housewares.

Though crystallinity up to 90 percent is achievable, high crystallinity makes the resin brittle. Its brittleness can be reduced, however, by copolymerization with small amounts of 1-butene or 1-hexene.[21] Copolymerization introduces structural irregularity without branching and reduces close packing and density. The increased segmental motion that results improves toughness and resistance to stress. These copolymers are the linear low density polyethylenes (LLDPE) (see above).

The greater toughness of LLDPE compared to HDPE is especially useful for higher-performance, low temperature applications. LLDPE is used widely for gas and water pipes and for other applications where a plastic needs to withstand high stress for long periods at moderate temperatures without failure. The properties of LLDPE are suitable for blown film, containers, and other packaging applications where film clarity is not important. It exhibits higher strength and puncture resistance, which allows the use of thinner films. About 3 billion lb of LLDPE are used annually for this purpose.

Polypropylene. Polypropylene and its copolymers account for about 13 percent of U.S. resin sales.[1] Its manufacture and properties have much in common with HDPE. Unlike polyethylene, however, isotactic, syndiotactic, Nomenclature").[22] Atactic polypropylene, and atactic configurations are possible as a consequence of the pendant methyl group (see above, under "Polymer Structure and

which can be produced by free radical polymerization, is primarily obtained as a by-product from the manufacture of the isotactic stereoisomer as described below. Since the random arrangement of the pendant methyl groups along the polymer chain prevents crystallization, the atactic material is an amorphous, low melting waxy solid with reduced solvent resistance. It is used as a water barrier coating for wire and cable, roofing material, and laminated paper. Syndiotactic polypropylene has been prepared in the laboratory, but no commercial interest has developed for it.[23]

Commercial polypropylene is based on catalysts of the Ziegler-Natta type that produce a product with an isotactic content of 90 percent or more. The presence of the methyl groups restricts movement of the polymer molecules somewhat, and crystallization rarely exceeds 65 to 70 percent. The degree of stereoregularity depends on a multitude of factors, some of which are not well understood. Variables include the type of metals in the catalyst, their oxidation states and crystal structure, and the organic functionality in the complex. In some systems both the rate of reaction and the degree of stereoregularity are increased by the addition of electron pair donors such as amines.

Isotactic polypropylene is a colorless thermoplastic that has a density of 0.91 to 0.94 g/cm^3 and is one of the lightest commercial plastics. Its chemical and physical properties are similar in many ways to those of HDPE. Although its impact strength is somewhat lower than that of HDPE, it is clearer, glossier, and harder. It also has a higher softening point (140–155°C, compared to 122–130°C for HDPE). It has unusually high resistance to flexing stress, and thus is used in molded articles with built-in hinges.

As with HDPE, copolymerization is used to modify the mechanical properties of isotactic polypropylene. Up to about 7 percent ethylene may be added to form a random copolymer. The ethylene in the chain reduces crystallinity somewhat, resulting in improved elongation and toughness. Although increasing the atactic to isotactic ratio can have a similar effect, copolymerization accomplishes it more efficiently. Terpolymers of polypropylene also are available.[24]

For applications requiring even higher impact strength, especially at low temperatures, further increases in random copolymer ethylene level become unsuitable because other important properties such as stiffness and melting point are degraded. For these applications ethylene is substituted for the propylene feed during the course of the polymerization, resulting in a block copolymer with up to 40 percent ethylene. The polymer consists of a block of highly isotactic polypropylene bonded to a block of random ethylene–propylene copolymer. At even higher ethylene levels, the copolymer takes on elastomeric properties. These copolymers are known as EPR or EPM rubbers.

The tertiary carbons along the backbone of polypropylene are susceptible to oxidation in air. This is especially a problem at high process or use temperatures, in sunlight, or during mechanical stress. Formulation with antioxidants, peroxide decomposers, and UV stabilizers greatly reduces this problem.

Polypropylene is used in a large variety of applications. About 20 percent of it is injection-molded into furniture, appliance parts, bottles, battery cases, luggage, distributor caps, screw-on bottle caps, and housewares. Up to 40 percent filler can be used to stiffen it. Various grades of filled polypropylene are used in industrial and automotive applications.

About 40 percent of polypropylene resin is extruded into fiber and filament, coatings, film, sheet, pipe, and wire insulation. The fiber does not absorb water and is stain- and abrasion-resistant. The largest market for polypropylene fiber is for carpet backing. The fiber also is used to make disposable diapers and other nonwoven fabrics. Because of its strength, lightness, and resistance to weathering and microorganisms, it has replaced natural fiber in rope and cord.

The strength and barrier properties of polypropylene film can be improved by stretching it during or just after extrusion, but before it has been cooled. This aligns the

polymer chains in the direction of drawing. The orientation is set during quenching, thereby increasing strength in the draw direction. Film stretched in one direction is called uniaxially oriented film. Film that is stretched in both the machine direction and the transverse direction is referred to as biaxially oriented. Oriented polymers are stressed and have lower entropy than unoriented polymers. When the unconfined oriented material is reheated to sufficiently high temperatures, the polymer strands randomize and the material changes dimensions. Stress relaxation is used to advantage in shrink wrapping.

The barrier properties of polypropylene film to certain gases such as oxygen can be improved further by laminating it to film with better barrier properties such as poly(vinylidene chloride). If one of the film components has a sufficiently low melting point, the laminate may be heat-sealable. Laminated barrier films are widely used for wrapping meats, bakery goods, snack items, and frozen foods. Form-and-fill packaging is the largest use of polypropylene film. The package is formed, filled with product, and sealed in a continuous process.

Vinyl Resins

Vinyl resins are the second largest class of industrial polymers, accounting for about 10 billion lb or 16 percent of domestic resin sales.[1] This class includes poly(vinyl chloride), poly(vinylidene fluoride), poly(vinyl acetate), poly(vinyl alcohol), the poly(vinyl ethers), and a numbet of other related polymers and copolymers.

Poly(vinyl chloride)

Poly(vinylidene chloride)

Poly(vinyl acetate)

Poly(vinyl alcohol)

Poly(vinyl methyl ether)

These are among the most versatile resins, based on the diversity of uses for them. Poly(vinyl chloride) and its copolymers are the most important of these resins.

Polyvinyl Chloride. Poly(vinyl chloride), or PVC, initially was used as a replacement for cellulose nitrate. It is prepared by the polymerization of vinyl chloride under free radical conditions, most of it by suspension polymerization and the rest by emulsion and mass polymerization. The uncompounded resin is an atactic, amorphous, rigid material. The large dipole moment of the carbon–chlorine bonds produces strong intermolecular associations that result in high melt viscosity and good dimensional stability at room temperature. The high chlorine content gives it a high degree of flame and chemical resistance, which makes it suitable for pipes and fittings, tank linings, and packaging for corrosives. Although PVC is insoluble in most hydrocarbons, it is partially soluble in polar solvents such as ketones and chlorinated hydrocarbons. Its melting point and glass transition temperature depend on the method of preparation. It softens between 70 and 80°C, and when heated above 80°C it begins to degrade, eliminating hydrogen chloride from the polymer backbone.

The heat and light sensitivity of PVC and its processibility are greatly improved when the resin is compounded with antioxidants, UV absorbers, lubricants, fillers, pigments, and plasticizers. Stabilization with antioxidants and UV screeners has permitted its use in demanding exterior applications such as water hoses, siding for buildings, and sheeting for roofs. PVC is unusual because it can accept very large quantities of plasticizer and still retain desirable properties over a wide temperature range. At high plasticizer concentrations the type of plasticizer strongly influences properties. Plasticizers can be used to tailor the product for a specific application. For example, citrate esters impart flexibility

for food contact applications; chlorinated hydrocarbons improve flame resistance and electrical properties; and the phthalic esters are lower-cost, general purpose plasticizers with good low-temperature properties.

The additives usually are mixed with the resin in two stages. Initially they are mixed at low shear to produce a good distribution. This is followed by roll milling or Banbury mixing, which homogenizes the mixture and forms a flux that is discharged to a mill for sheeting, or to an extruder for shaping. The molten sheet then can be calendered, compression-molded, or cooled and pelletized for packaging, storage, and shipment prior to fabrication.

Over 60 percent of the PVC resin produced is extruded, much of it into pipe and conduit for water, gas, drainage, and sewerage. A large amount is extruded into siding, window profiles, window blinds, gutters, and downspouts. Calendered PVC is made into film, sheet, flooring, paneling, pool liners, upholstery, shower curtains, wall covering, electrical tape, packaging, and labels. PVC also is blow-molded into bottles. Laminates of PVC and cloth are used for upholstery fabrics, leather substitutes, and interior trim in domestic and automobile use. Because of its good electrical properties, stability, and toughness, PVC is used widely for both domestic and industrial wire and cable sheathing. Adding various-colored pigments aids in the identification of complex circuits.

PVC can be blended with other polymers to produce injection-molding compounds. Properties that make these blends suitable for appliance and business equipment housings such as computer monitors and keyboards include good impact strength, heat resistance, flame retardancy, and chemical and solvent resistance.[25]

Other important forms of PVC include plastisols, which are dispersions of the polymer in a plasticizer, and organosols, dispersions of polymer, plasticizer, and a volatile solvent. Dispersions are produced by spray-drying the polymer mixture made by emulsion or suspension polymerization. The friable powder produced this way is milled and then mixed with plasticizer and other additives. At ambient temperatures these mixtures are low-viscosity, easily handled liquids. When these dispersions are heated to temperatures above 125°C, any volatile solvents present evaporate, and the polymer and plasticizer fuse into a homogeneous solid. Organosols and plastisols are made into products by painting, dipping, molding, casting, and foaming. They are used in the production of handles and grips, plastic cloth, belting, film, coatings, sheathing, and a variety of molded and cast products.

Other Vinyl Resins. One of the most important uses of poly(vinylidene chloride), PVDC, is for packaging film. The PVDC resin used for this purpose is usually a copolymer of vinylidene chloride and vinyl chloride. The film is produced by casting or by blown film extrusion. Biaxially oriented film is clear, with good strength and low permeability to moisture and other gases. This makes it especially useful for food packaging.

Homopolymerized poly(vinyl acetate) (PVA) usually is made by an emulsion process.

Vinyl acetate Poly(vinyl acetate)

The solid resins are clear, colorless, odorless, and tasteless, with a wide molecular weight distribution and a broad softening range. They are soluble in ketone and ester solvents and are readily plasticized. PVA is used in coatings and adhesives, as a chewing gum base, and as an intermediate in the production of poly(vinyl alcohol), PVOH.

PVOH is made by basic hydrolysis of PVA.

Poly(vinyl alcohol)

Its solubility in water depends on how it is prepared. It is used for packaging film, as a thickening agent, as a component in adhesives, in the manufacture of controlled-dosage film sachets, and for the production of poly(vinyl butyral) (PVB).

PVB is made from hydrolyzed PVA by reaction with butyraldehyde.

Poly(vinyl butyral)

It has excellent optical properties, and, when plasticized, has good toughness and adhesion to glass. The largest use of plasticized PVB is for a safety glass inner layer. Its properties are adjusted to cushion the impact of an automobile occupant with the windshield during an accident. It also adheres to shards of glass broken during a collision or when a stone hits the windshield.

A variety of resin products are prepared from copolymers of ethylene and vinyl acetate. Though the proportion of ethylene can vary greatly, the most important ethylene vinyl acetate (EVA) copolymers contain less than 50 percent vinyl acetate. EVA copolymers with low vinyl acetate content are more crystalline than those with higher vinyl acetate content. In general, the more amorphous copolymers have higher impact strength, greater elasticity, and better resistance to stress cracking than LDPE. EVAs find use in packaging film and in a variety of adhesive products.

The hydrolyzed derivative of EVA is ethylene vinyl alcohol copolymer (EVOH). EVOH resins have high strength and elasticity. The films are glossy and clear and have good oxygen, nitrogen, and carbon dioxide barrier characteristics. Much EVOH is used for food packaging. Because the gas barrier properties of EVOH are reduced under humid conditions, it often is coextruded with nonpolar resins such as polyethylene or polypropylene to form multicomponent laminates.

For packaging applications that demand it, the moisture and gas barrier properties of the laminated film are further enhanced by metallizing one of the components. Because a continuous metal layer is much less porous than plastic, the rate of diffusion of gases thus is significantly reduced. The metallizing process is carried out in a high vacuum chamber where a very thin coating of aluminum is evaporated onto the surface of the plastic film. The development of multilayer films with exceptional barrier properties has made packaging of foods with nitrogen gas possible. This has substantially improved the quality and shelf life of many packaged food products.

Polystyrene and Styrene Copolymers

Polystyrene and its copolymers account for about 10 percent of the domestic resin market.[1] In its unmodified state, polystyrene is a hard, crystal-clear thermoplastic with relatively poor impact strength. Styrene is one of only a few vinyl monomers that can be polymerized by free radical, cationic, anionic, and Ziegler-Natta processes.

Styrene Polystyrene

When prepared by free radical and ionic processes, it is an atactic, amorphous solid. Isotactic polystyrene has been prepared by Ziegler-Natta polymerization.[20] It is an opaque, semicrystalline material that has a higher T_g and melting point and lower solubility than the more common amorphous atactic stereomer. The syndiotactic form also has been made.[26] However, no commercial interest has yet been shown for the stereoregular materials.

Amorphous atactic polystyrene softens at 90 to 95°C and becomes a mobile liquid at 140°C. It has good electrical properties and good resistance to strong acids and bases. It is soluble in esters and aromatic and chlorinated hydrocarbons and insoluble in aliphatic hydrocarbons and lower alcohols.

A large number of grades of polystyrene have been made for specific purposes, most of them by manipulating molecular weight and molecular weight distribution and by adding lubricants and plasticizers. The resins can be injection-molded, extruded into sheet and film, and made into foam. The electrical and heat-resistant grades are used as electrical insulation for capacitors and transformers and as housings for electrical appliances. Other grades are used for packaging, jars and tumblers, combs and hairbrushes, coat hangers, food canisters, and tubs. Foamed polystyrene is used as thermal insulation for refrigerators, coolers, trays, and cups; for mechanical shock insulation in molded packaging for cameras, small appliances, and radios; and for flotation in life preservers, buoys, and docks.

Toughened or rubber modified polystyrene blends are prepared by incorporating up to 10 percent by weight of polybutadiene or styrene–butadiene copolymer rubbers into the resin. The dispersed rubber particles decrease the plastic's brittleness by interrupting the crack propagation process. Unlike the unblended polymer, rubber modified polystyrene is translucent. The volume of rubber modified polystyrene made is roughly the same as that of general purpose polystyrene.

Good adhesion between the dispersed rubber phase and the continuous polystyrene phase is one requirement for improving the impact strength of rubber modified blends. Adhesion can be improved by compatibilization. There are several ways that this can be done. One way is to add molecules to the resin mixture that interact with, or have molecular similarities to, both phases. These are called compatibilizers. Another way to improve adhesion is to graft reactive substituents onto the molecules of one or both phases. Properly chosen, these grafted units interact with molecules of the other phase to improve the adhesive bond strength.

Rubbers are incorporated by mechanical blending and by polymerizing styrene monomer containing dissolved rubber. The latter process usually results in better impact strength due to efficient grafting between the growing polystyrene chains and the rubber. Strength and toughness are closely tied to the particle size of the rubber phase. Rubber-modified products made this way often are referred to as high impact polystyrene (HIPS).

Some HIPS plastics are extruded into sheet for subsequent thermoforming, which is done by heating the sheet above its softening point and forcing it against a mold by applying vacuum, air, or mechanical pressure (see below, "Thermoforming" in section on "Fabrication of Plastics"). Thermoforming of HIPS sheet for refrigerator door liners was one of the developments that helped to promote the growth of rubber-modified materials. HIPS is used in a variety of other products, including dinnerware, dairy containers, cold drink cups, and camper/trailer covers. In packaging applications, HIPS frequently is coextruded with a vinylidene chloride copolymer. The HIPS component provides structural strength and thermoformability, and the vinylidene chloride copolymer provides good gas barrier properties. In some packaging applications, such as yogurt containers, HIPS is coextruded with clear polystyrene, which adds a glossy surface that improves the package's visual appeal.

Styrene Copolymers. Copolymerization is another way to improve the mechanical properties and chemical resistance of polystyrene. Acrylonitrile, butadiene, alpha-methylstyrene, methyl methacrylate, divinylbenzene, maleic anhydride, and other monomers have been copolymerized with styrene to produce commercially important copolymers. Some of the most widely used of these are those prepared with acrylonitrile and butadiene. Styrene copolymerized with butadiene (SBR) is one of the more important elastomeric materials used today. (See Chapter 18.)

Styrene–acrylonitrile (SAN) is produced by mass, suspension, and emulsion polymerization. The copolymer has better chemical and mar resistance than the homopolymer. Properties vary with the ratio of the monomers; in general, an increase in the acrylonitrile content improves toughness and chemical resistance, but also increases color and melt viscosity. SAN is used for cups, handles, housings, and appliance parts. Low-color resins with high light transmittance are now available. Good optical properties make SAN useful for tumblers, serving dishes, canisters, and some disposable medical products.[27]

Acrylonitrile–butadiene–styrene (ABS) blends have become an important class of structural polymers. Like HIPS, ABS is a two-phase system in which the rubber is dispersed in the rigid SAN copolymer matrix. It is prepared by grafting SAN onto polybutadiene or butadiene copolymer. Properties can be modified by altering the ratio of the three ingredients and by manipulating rubber particle size, morphology, and the degree of grafting. ABS has high impact strength, toughness, and hardness and better resistance to outdoor weathering conditions than HIPS and SAN. Added protection from ultraviolet radiation can be obtained by adding pigments, stabilizers, and fillers.

ABS has good melt strength and little tendency to orient or develop mechanical anisotropy during molding. It can be made into uniformly tough molded and thermoformed objects. Moldings and extrusions with high gloss can be painted, metallized, and electroplated. ABS is extruded into pipes for plumbing and liquid transportation. Its high strength and heat and chemical resistance make it suitable for TV, telephone, computer, and camera housings, and for refrigerator linings, safety helmets, fascia, and dashboards in autos and many other applications.

Acrylics

Acrylic resins are prepared by addition polymerization from acrylic or methacrylic acids and their ester, nitrile, and amide derivatives.

R=H Arylic acid
R=CH_3 Methyl acrylate

R=H Methacrylic acid
R=CH_3 Methyl methacrylate

Acrylonitrile

Acetonitrile

They are produced by a variety of processes including bulk, solution, emulsion, and suspension polymerization.

Because of the electronic effect of the carbonyl group, acrylates and methacrylates readily take part in both free radical and anionic addition polymerization reactions. The type of substituent attached to the double bond significantly affects polymer properties, and because these monomers readily undergo copolymerization, considerable property diversity can be designed into the polymers that are formed.

Polymers derived from acrylic and methacrylic acid esters are noted for waterlike clarity and color and for excellent stability during aging. Depending on the nature of the ester group and polymerization conditions, the polymer can be tacky, rubbery, or glasslike.

Thus, acrylics are used for such diverse products as pressure-sensitive and structural adhesives, coatings, and rigid plastics. Because of their low T_g, poly(ethyl acrylate) and poly(butyl acrylate) are used in acrylate-based rubbers.

The most important acrylic plastic is polymethyl methacrylate (PMMA). It also was the first acrylic to be produced commercially. It is an amorphous thermoplastic polymer that can be thermoformed, molded, or extruded above 150°C. It is soluble in ketones, esters, and chlorinated hydrocarbons, but is resistant to attack by dilute acids and bases. Compared to polymethyl acrylate, PMMA is more stable to exterior aging, has lower moisture absorption, and is more resistant to alkalies.

Applications for PMMA take advantage of its combination of weather resistance, clarity, light weight, impact strength, hardness, and formability. It has a visible light transmission of about 92 percent, which is somewhat higher than that of glass. Consequently it is used as a substitute for glass in some applications, including bank teller windows, skylights, archways, exterior signs, corrugated roofing, storm doors, and bath and shower enclosures. It also is used for interior applications such as appliance housings, tumblers, and instrument panels. One of its major disadvantages compared to glass is low abrasion and solvent resistance. For applications where abrasion and solvent resistance are important, PMMA sheet is available with a thin hard coating that improves its durability considerably.

The high degree of intermolecular interaction between the polar nitrile groups in poly(acrylonitrile) restricts bond rotation and produces a stiff polymer. As a result, poly(acrylonitrile) has a high melting point and is soluble in only a few solvents. It cannot be melt-processed because of extensive decomposition. However, copolymerizing acrylonitrile with other monomers produces a more easily processible material. Much of the commercial output of acrylonitrile monomer is used for fiber and for structural plastics such as SAN and ABS.

Some of the properties of polyacrylic acid, polymethyacrylic acid, and polyacrylamide are similar. They are hard, brittle, transparent, and hygroscopic solids; they are soluble in water and polar organic solvents, and insoluble in nonpolar solvents. They are used as thickeners, binders, adhesives, and suspending and flocculating agents, but the volume used is comparatively small. Several applications for polyacrylamide and partially hydrolyzed derivatives rely on their ability to flocculate solids in aqueous suspensions. Thus, they are used for water purification, especially in paper making and mining operations. They also are used to improve the flow properties of water, to adjust the dry strength of paper, and in secondary and tertiary oil recovery.

Polyesters

Polyesters are a class of resins with repeating ester linkages in the polymer backbone. Nominally, they are derived from the condensation of dicarboxylic acid monomers with diol monomers, or by self-condensation of hydroxycarboxylic acids.

or

Though a variety of aliphatic polyesters are available, the major commercial thermoplastic polyesters are based on the aromatic dicarboxylic acid, terephthalic acid. The terephthalic acid polyesters prepared from the lower glycols are tough crystalline solids. The isophthalic acid derivatives have lower T_g's

than other members of the group and are amorphous, or crystallize with difficulty. This illustrates the importance that molecular symmetry in the repeating units has on the physical characteristics of polymeric materials.

Poly(ethylene terephthalate) (PET) the most important of this group, was commercialized in the early 1950s, and has found a broad range of uses. The largest uses are in fiber staple; photographic film; video, magnetic, and computer tape; electrical insulation; bottles; and engineered products. Its advantages for these applications include its high melting point, low water absorption, good surface gloss, dimensional stability, and good control over properties.

PET can be prepared in several ways. In a widely used batch procedure, dimethyl terephthalate is mixed with excess glycol and a basic catalyst, and the mixture is heated to temperatures around 200°C. Methanol and excess glycol are removed to produce the intermediate bis-(2-hydroxyethyl) terephthalate. In a second stage, the temperature is increased to near 300°C under a slight vacuum to carry off residual volatiles produced in the final condensation step. The reaction is terminated when the viscosity increases to a predetermined value. The product is cooled and converted into granules for packaging and shipping.

Resin with an average molecular weight of 35,000 to 40,000 is used for film and fiber. Fiber is made from the resin by spinning, and film is manufactured by extrusion. Both processes orient the polymer molecules in the direction of drawing. Film is generally also stretched in the cross-machine direction to make a biaxially oriented product. Such films have improved strength and modulus as well as better optical properties. The hot film is quenched after annealing under tension for a predetermined period of time to allow further crystallization to take place. This yields a product with controlled clarity, strength, and modulus. Orientation also will improve heat resistance and barrier properties. Once the material has been fine-tuned to this degree, it becomes temperature-sensitive. Heating to near the T_g makes the polymer strands more mobile and causes additional crystallinity, loss of clarity, and dimensional changes. When this is done intentionally to stabilize a plastic, it is called annealing.

Resin with a molecular weight of about 70,000 is used for molded parts. Although the low rate of crystallization is helpful in making film, it is a problem for making castings and injection-molded engineered products. However, when the melt is allowed to crystallize under controlled temperature conditions in the presence of finely divided inorganic nucleating agents, thick materials can be produced with the desired morphology.

When tetramethylene glycol is used in place of ethylene glycol, poly(butylene terephthalate) (PBT), a material with a lower T_g and melting point (220–235°C), is obtained. Because PBT crystallizes more rapidly than unmodified PET, its mold cycle times are faster. It also has good strength and rigidity and has found use in precision parts for electronic devices, office appliances, and automobiles. Its material properties can be further modified with a variety of additives. The inclusion of glass fiber increases its modulus and its heat deflection temperature.

PET and PBT are more expensive than commodity plastics, but they have many advantages in structural applications, including rigidity, surface hardness, machinability, and higher service temperatures. Though they have been mentioned here, structural grades of the polyesters can be classified as engineering thermoplastics.

Engineering Thermoplastics

Engineering plastics have been defined as "thermoplastics that maintain dimensional stability and most mechanical properties above 100°C and below 0°C."[28] Compared to metal, wood, glass, and ceramic materials that they might replace, engineering plastics offer the advantages of corrosion resistance, weight reduction, and fabrication ease. For example, a large market for engineering plastics is developing in the transportation industry where weight reduction, impact resistance, design flexibility, corrosion resistance, and cost savings from parts consolidation are incentives for materials to replace steel.

An engineering plastic is expected to function when it is subject to high mechanical stress, temperature extremes, and hostile environments. In contrast, commodity thermoplastics such as the polyolefins, polystyrene, and PVC lose important mechanical properties when temperatures approach 100°C.

The engineering plastics group includes the polyamides, polyimides, polyacetals, polycarbonates, polyethers, polysufones, polysulfides, fluoropolymers, some of the polyesters, some grades of ABS, and the so-called high performance plastics. Together they account for about 5 percent of the plastics market. The higher-performance segments of this class cost more than conventional engineering plastics, but they retain mechanical and electrical properties during continuous use at temperatures above 150°C. Table 19.3 lists some of the major representatives of this group, the relative quantities sold, and approximate continuous use temperatures.[29]

Some of these resins, such as the polyacetals, nylons, and polyether ether ketones, are highly crystalline. The crystalline resins are more resistant to solvents and grease and are used for parts such as gears, bearings, and petroleum product containers, and parts that need resistance to aggressive solvents and grease. However, in applications where clarity and toughness are important, such as security glazing, clear piping, or protective housings, an amorphous material such as polycarbonate might be used. The crystalline materials tend to be more dimensionally stable under stress

TABLE 19.3 Engineering Thermoplastics

	Approximate Continuous Use Temperature (°C)	*% of Total*
Polyamides	80	36
Polycarbonates	120	27
Polyacetals	80	16
Polyphenylene oxide	90	10
PET/PEB	130	9
High performance resins		2
Polyphenylene sulfide	200	
Polysulfone	150	
Polyether sulfone	180	
Polyether ketone	250	
Polyether ether ketone	240	
Polyetherimide	170	

but shrink more in molding operations than the amorphous ones.

Engineering thermoplastics can be fabricated by conventional means into mechanically functional parts or structural components. They often are modified by adding stabilizers, fire retardants, fillers, and reinforcement materials. Only a few members of this diverse class of resins will be discussed here.

Polyamides and Polyimides. The polyamides are a class of polymers characterized by repeating amide links in the polymer backbone:

They include natural polymers such as silk and synthetics such as nylons. Nylon polymers first were synthesized by Wallace Carothers and coworkers at Du Point in the late 1920s and 1930s. They are derived from carboxylic acid and amine precursors by both condensation and ring opening polymerization. They are commonly named by adding to the word nylon, a number equal to the number of carbons in the parent compounds. Thus, nylon 6.6 is the product of the condensation reaction between hexamethylenediamine, $NH_2(CH_2)_6NH_2$, or HMDA, and adipic acid or its acid halide and methyl ester.

Nylons also can be made by the self-condensation of amine-substituted carboxylic acids. Nylons 7, 9, and 11 are made this way.

Nylons also are prepared by the ring-opening addition polymerization of certain lactams. The most important of these ring-opening processes is the polymerization of caprolactam to give nylon 6.

Of the many nylons available, nylons 6.6, 6, 11, and 12 are the most widely used. Nylons prepared from the terephthalic acids have very high melting points and are called aramides. The aramides are used in fiber applications where high thermal stability and high tensile strength are important. The repeating unit (R) can be derived from aliphatic or aromatic diamines.

Most nylons are highly crystalline and undergo a sharp transition from a tough, rigid solid to a low-viscosity melt. This makes them easy to process. Nylons having an even number of carbon atoms between the amide groups pack better and therefore have a higher melting point than comparable nylons with odd numbers of carbon atoms. The crystalline nylons are translucent and have good chemical and hydrolytic stability.

Though the nylons originally were developed for their fiber-forming characteristics, their first commercial application was for toothbrush bristles, and shortly thereafter they were used for women's stockings. Not only are they drawn into monofilament and spun into fiber; they are extruded into film and tubing, blow-molded into bottles, injection-molded

into numerous products, and used in coatings and adhesives. Large quantities of nylon are used for surgical suture, rope, tire cord, carpet stable, and brush bristles. Their advantages for these applications include durability, tensile strength, toughness, and flexibility.

The nylons were among the first engineering thermoplastics produced on a large scale. In these applications use is made of their toughness, high melting point, chemical resistance, and self-lubricating properties. When nylons are fiber-reinforced, their heat deflection temperatures approach the crystalline melting point, and because they retain toughness and flexibility at low temperatures, they can be used over a wide temperature range. Examples include pipes, gears, bearings, cams, housings, impeller blades, casters, and door catches.

Nylons can reversibly absorb up to about 10 percent moisture. The moisture uptake changes dimensions and mechanical properties. Nylons 11 and 12, which have a higher proportion of hydrophobic methylene groups than nylon 6, for example, have a lower moisture uptake. Nylons can be mineral-filled or reinforced with fiber. Modifications of this kind significantly decrease water uptake and increase modulus, impact strength, temperature stability, and fire retardancy. Heat distortion temperature and creep also are improved this way. Filled nylons have replaced metals in applications such as door handles, hinges, window and door fasteners, and oil dipsticks.

Related to the polyamides are the polyimides, which contain the characteristics imide linkages in the backbone:

Although polyimides can have an open-chain structure, cyclic polyimides have found more use because they are more thermally stable than open-chain polyimides.[30] The polyetherimide prepared from pyromellitic dianhydride and 4,4′-diaminodiphenyl ether has found use in high temperature coatings, adhesives, and structural plastics. The reaction involves rapid formation of the polyamide, called polyamic acid, followed by a high-temperature ring closure step.

Thermodynamic factors favor ring formation over competing cross-linking reactions. These steps usually are carried out in stages, with ring closure driven to completion at higher temperatures. The polyamic acid produced during the first stage can be applied as a melt or in some cases as a solution, in a solvent such as dimethylformamide or dimethyl sulfoxide. This varnish can be used to impregnate glass fabrics or to bond composites. Once it is applied to the substrate, the temperature is increased to drive off solvents and complete ring closure. The final article then may be finished by machining.

Many types of polyimides are possible, although those derived from aromatic precursors are most important because they are the most resistant to hydrolysis and have excellent thermal and oxidative stability at temperatures up to 300°C or more. The resistance of polyimides to abrasion and solvents is good, and filled grades are available for load-bearing applications. Polyimide foams can be made by the reaction of dianhydrides with diisocyanates. The carbon dioxide that is produced acts as a blowing agent.[31]

Polyimides are good insulators and particularly resistant to electrical and radiation damage. They are used as structural adhesives, for electric motor insulation, for high performance cable and wire coatings, and for printed circuit boards.

Polycarbonates. Polycarbonates are polyesters derived from carbonic acid and have the backbone structure:

Although aliphatic polycarbonates have been produced, the term "polycarbonate" usually refers to the bisphenol A esters and copolymers of carbonic acid. Bisphenol A polycarbonate has been made by transesterifying diphenyl carbonate with bisphenol A, but most now is manufactured by interfacial polymerization of bisphenol A with phosgene.

In the interfacial process, bisphenol A is dissolved in the aqueous sodium hydroxide phase of a two-phase system with methylene chloride. As phosgene is pumped into the mixture, the sodium chloride condensation by-product dissolves in the aqueous phase, and the growing polymer dissolves in the organic phase. A tertiary amine is used as the catalyst. When the reaction is complete, the methylene chloride solution is extracted with acid to remove basic components and then washed with water. The solvent is then flashed off and recycled, leaving the solid.

It is an amorphous material that exhibits unusually high impact strength over a broad temperature range. Though its T_g approaches 150°C, it can exhibit ductile failure at temperatures as low as −200°C. This is unusual for an amorphous material, as most thermoplastics are brittle below T_g.

In its fused form polycarbonate has glasslike clarity with 89 percent transmission of visible light, high thermal resistance, and good processibility. Material with an average molecular weight of 22,000 to 50,000 softens above 220°C. It can be injection-molded and extruded between 270 and 320°C at pressures between 10,000 and 20,000 psi. The resin must be dried prior to processing in order to prevent thermally induced hydrolysis.

Polycarbonate is resistant to dilute acids, aliphatic hydrocarbons, alcohols, and oils, but is attacked by alkalis, aromatic and chlorinated hydrocarbons, ketones, and esters. Methylene chloride and pyridine are good solvents. Scratches and crazing by solvents and chemicals substantially lower its impact strength. For applications where this can be a problem, sheet is available with an abrasion-resistant coating that significantly improves resistance to abrasion and solvents. Polycarbonate cannot be used for extended periods outdoors without UV-absorbing coatings because photodegradation reduces its strength and causes it to discolor.

A large fraction of the polycarbonate made is used by the electronics industry. It has electrical insulating properties that are nearly independent of environmental conditions. This includes high surface and volume resistivity and good dielectric properties. Capacitors are made from metallized grades. A large outlet for high grade polycarbonate is compact disks. The transparency and impact resistance of polycarbonate has led to its use for motorcycle windshields, high-speed aircraft windows, ophthalmic lenses, helmets, and security glazings for banks, offices, and armored vehicles. It also is used in medical appliances, auto head- and tail-light housings, fan blades, and machine parts. It is extruded into film for packaging and pressure-activated switch protectors.

Both random and block copolymers of polycarbonate have been made. Copolymers with terephthalic and isophthalic acid units are used in glazing applications. Polymerization with mono- or difunctional poly(dimethylene) oxide gives, respectively, ABA and $(AB)_n$ block copolymers with improved heat distortion temperature. Blends with ABS exhibit good toughness and electrical properties. Polycarbonate–ABS blends are used for dashboard panels, spoilers, wheel caps, housings for lamps, and office equipment.

Polyetherether Ketones, Polysulfones, and Acetal Resins. The poly(etherether ketones), or PEEK, can be prepared by condensation of the dipotassium salt of hydroquinone or bisphenol A with 4,4′-difluorodiphenyl ketone at temperatures in excess of 300°C.

Hydroquinone dipotassium salt — 4,4′-Difluorodiphenyl ketone

ether link — ketone link

They are crystalline resins with T_g's in excess of 140°C and melting points of around 320°C. Their heat deflection temperature can be increased to near their melting point by incorporating glass filler. They have excellent solvent resistance and stability at 200°C and can be heated to near 300°C for short periods with no serious effect on their properties. Like other resins with the photosensitive aromatic backbone, they are not suitable for extended use outdoors.

PEEK can be processed into filament, tubes, sheet, and profiles by conventional methods at melt temperatures near 370°C. The largest user of the resin is the electronics industry, which uses it for cable insulation, connectors, and flexible circuit boards. It also is used for valve seals and compressor plates.

Polysulfone resins are prepared by a process similar to that used to prepare polyether ketones, except that the displacement occurs on a halogen-substituted aromatic sulfone.

They are amorphous, yellow, transparent, rigid resins with T_g's over 180°C and very good resistance to thermal oxidation. Like PEEK, they have good electrical properties. They are inherently resistant to burning, and for most applications do not require fire retardants. They also can be used in applications requiring long exposure to steam or hot water. Because they are transparent to microwaves, they are suitable for microwave cookware. Other uses include appliance parts, gas separators, medical devices, and electrical components.

The acetals, which have the $(—CH_2O—)_n$ ether repeating unit, are prepared by ionic polymerization of formaldehyde and its oligomers. Initiators include alkoxide, amine, and phosphine bases, and both mineral and Lewis acids.

Initiation by bases, B^-:

$$B^- + CH_2{=}O \rightarrow B—CH_2—O^-$$

$$B—CH_2—O^- + n\,CH_2{=}O \rightarrow B—CH_2—(O—CH_2)_n—O^-$$

Initiation by acids, A^+:

$$A^+ + CH_2{=}O \rightarrow A—OCH_2^+$$

$$A—OCH_2^+ + n\,O{=}CH_2 \rightarrow A—O—(CH_2—O)_n—CH_2^+$$

Up to 15 percent ethylene oxide may be used as a comonomer to provide flexibility and stability. At one or both chain ends is a hydrolytically unstable $—OCH_2OH$ hemiacetal group. In moist environments or at elevated temperatures, formaldehyde is split off, generating another hemiacetal group.

$$\text{-}(CH_2O\text{-})_n\underbrace{CH_2—OH}_{\text{unstable hemiacetal end group}} \xrightarrow{\text{heat}} \text{-}(CH_2O\text{-})_{n-1}CH_2OH + CH_2{=}O$$

This unzipping, which is essentially the reverse of the polymerizing reaction, can be prevented by esterifying or alkylating the hydroxyl end group.

$$\text{-}(CH_2O\text{-})_n\underbrace{CH_2—OC(=O)CH_3}_{\text{stable acetate ester end group}}$$

The acetals are translucent, crystalline materials with good solvent resistance. Parts made from them exhibit high stiffness, good tensile strength, and dimensional stability. Like the nylons with which they compete, they are self-lubricating, but unlike the nylons, they retain dimensional stability under ambient moisture conditions. They have good fatigue resistance, but relatively low impact resistance. As in other polymeric systems, their impact resistance is increased significantly when they are reinforced with fibers.

The melting points of the homopolymer and ethylene oxide copolymer are around 175°C and 165°C, respectively. Typical of highly crystalline resins, liquid flow temperatures are about 10°C higher. They are injection-molded and extruded. Their utility is based on lower cost and improved performance in applications that were formerly met by zinc, brass, aluminum, and cast steel parts. Examples include fuel pump and filter housings, fuel caps, bearings, gears, cranks, springs, pens, integral hinges, zippers, and solvent containers.

IMPORTANT CLASSES OF PLASTICS AND RESINS: THERMOSETS

Cross-linking makes thermosets infusible and insoluble. The cross-linked network can form during polymerization when multifunctional monomers are present. Thermosets also are formed when resins with reactive functionality are caused to react with cross-linking co-reactants. Alternately, a cross-linked network will be produced when thermoplastic resins are irradiated or vulcanized or undergo autoxidation.

Thermoset plastics can be cross-linked to various degrees; they can be soft, flexible, elastomeric materials, or hard and structurally rigid. Though thermosets account for only about 15 percent of the plastics market, they are an extremely important segment.[1] The most important thermosets are derived from polyurethane, phenolic, unsaturated polyester, and epoxy resins.[32]

Polyurethanes

The polyurethanes, a versatile class of polymers that were discovered by Otto Bayer and coworkers in 1937,[33] are characterized by repeating carbamate units (—OCONH—) in the polymer backbone. They generally possess good toughness and abrasion and chemical resistance, and are made into film, sheet, foam, fiber, coatings, adhesives, and sealants. They also find extensive use in mechanical, sound, and thermal insulation, and for medical applications such as catheters and artificial blood vessels.[34]

Polyurethanes are prepared by the addition of polyols to polyfunctional isocyanates.

$$n\ \mathrm{HO{-}R{-}OH} + n\ \mathrm{OCN{-}R'{-}NCO} \longrightarrow \left(\mathrm{O{-}R{-}O\overset{\overset{\displaystyle O}{\|}}{C}NH{-}R'{-}N\overset{\overset{\displaystyle O}{\|}}{C}}\right)_n$$

Polyurethane backbone

Commonly used isocyanates include:

Toluene diisocyanate (TDI)
(a mixture of 2,4- and 2,6-isomers)

4,4′-Diphenylmethane diisocyanate (MDI)

$$\mathrm{OCN{-}(CH_2)_6{-}NCO}$$

Hexamethylene diisoyanate (HDI)

Dicyclohexylmethane diisocyanate (DCHDI)

Because the functionality, equivalent weight, and backbone structure of the reactants can be varied over a wide range, a broad spectrum of properties can be produced. The products will be thermoplastic if the functionality of the monomeric ingredients are two or less, or thermoset if functionality is greater than two. By far the largest class of industrial polyurethanes are thermosets derived from aromatic isocyanates.

In addition to alcohols, isocyanates can react with other substances with polar hydro-

gens. Examples include:

1° and 2° amines $R{-}NCO + R'{-}NH_2 \longrightarrow R{-}NH\overset{\overset{\displaystyle O}{\|}}{C}NH{-}R'$

A urea

Carboxylic acids $R{-}NCO + R'{-}\overset{\overset{\displaystyle O}{\|}}{C}OH \longrightarrow R{-}NH\overset{\overset{\displaystyle O}{\|}}{C}O\overset{\overset{\displaystyle O}{\|}}{C}{-}R'$

An anhydride

Ureas $R{-}NCO + R'{-}NH\overset{\overset{\displaystyle O}{\|}}{C}NH{-}R'' \longrightarrow R'{-}N\overset{\overset{\displaystyle O}{\|}}{C}NH{-}R''$ with $\underset{\underset{\displaystyle NH{-}R}{|}}{C{=}O}$

A biuret

Urethanes $R{-}NCO + R'{-}NH\overset{\overset{\displaystyle O}{\|}}{C}O{-}R'' \longrightarrow R'{-}N\overset{\overset{\displaystyle O}{\|}}{C}O{-}R''$ with $\underset{\underset{\displaystyle NH{-}R}{|}}{C{=}O}$

An allophanate

Water $R{-}NCO + H_2O \longrightarrow R{-}NH\overset{\overset{\displaystyle O}{\|}}{C}OH \longrightarrow R{-}NH_2 + CO_2$

A carbamic acid

Hence, care must be exercised in the choice of reagents, solvents, and reaction conditions to prevent unwanted side reactions during the production of polyurethanes. Reagents must be dry and free of unwanted alcohols. The relative rates at which the various isocyanate reactions proceed are as follows: aromatic isocyanates > aliphatic isocyanates. Also: aliphatic amines > aromatic amines > primary alcohols > water > secondary alcohols > tertiary alcohols > phenols > carboxylic acids > urethanes. These reactions are catalyzed by tertiary amines and a variety of metal salts. The most active catalysts are organo-tin compounds such as dibutyl tin dilaurate.[35]

An important fabrication process that takes advantage of the high reactivity of the catalyzed isocyanate–polyol reaction is called reaction injection molding (RIM). In RIM, precise amounts of the reactive components are metered to a mixing head, where the liquid streams are mixed and injected into a mold. Bulky, complex shapes and thin-walled articles can be produced in this way. Reaction times of a few seconds are possible, and mold residence times of one or two minutes have been achieved. RIM consumes less energy and uses less expensive equipment than conventional injection-molding processes. The process is used to make fascias, bumpers, and other flexible automotive parts.[36] Fillers may be incorporated in the components to lower cost and increase rigidity. This fabrication technique is discussed further in the section on "Fabrication of Plastics" under "Injection Molding."

Polyurethane foams are produced by synthesizing the polymer concurrently with gas evolution from a blowing agent. Bubbles or gas are trapped in the reacting mixture to form a cellular structure. Commercial nitrogen-forming blowing agents include azobisformamide and azobisisobutyronitrile. In some processes water is deliberately added to

generate carbon dioxide. The amine by-product produced this way reacts further with unreacted isocyanate to generate biuret cross-links. In rigid foams, low-boiling halogenated alkanes such as trichlorofluoromethane have been used. Foams produced this way are very good insulators because the halocarbon blowing agents trapped in the cells have low thermal conductivity. However, concern about ozone depletion is causing the elimination of many halocarbon blowing agents.

Foams usually are made with a mixture of aliphatic polyether or polyester diols and triols. A typical formulation also will contain an aromatic diisocyanate such as toluene diisocyanate or diphenylmethane diisocyanate, a catalyst, surfactants, and foam stabilizers. The inclusion of trifunctional reactants makes varying degrees of cross-linking possible. Lightly cross-linked foams are flexible, whereas those more highly cross-linked are rigid. Flexible foams generally have an open cell structure; most rigid foams are made with closed cells.

Polyurethane foams have dry heat stability to about 150°C and are resistant to many solvents. Because more than 90 percent of a low density foam consists of gas, these materials exhibit low thermal conductivity and are good acoustical insulators. Compared to polyether foams, polyester foams are stiffer and more resistant to dry-cleaning fluids. Thus polyester foams are preferred in such applications as textile laminates and shoulder pads. Polyether polyols produce a foam that is more resilient and less subject to hydrolysis than other foams; so polyether foams are preferred for upholstery, bedding, and carpet backing where cushioning is important. Rigid foams are used for lightweight insulated construction panels, packaging, and flotation equipment. Further information is included in the section on "Fabrication of Plastics" under "Foamed Plastics."

Phenolic Resins

The phenolics were the first wholly synthetic plastic materials commercialized. The most important phenolic plastics are derived from the condensation of formaldehyde with phenol or its derivatives. Though this class of polymer was first made in the 1870s by von Baeyer, Leo Baekeland found early in the twentieth century that useful moldings can be made if the final stages of the curing reaction are carried out under heat and pressure.

The chemistry of the phenolics is complex. Polymerization involves heating the reactants with acidic or basic catalysts until condensation begins. Depending on the ratio of reactants and the catalyst type, materials with different properties are formed. If the molar ratio of formaldehyde to phenol is 1:1 or greater, and a basic catalyst such as sodium or calcium hydroxide is used, then a reactive resin called a resole is formed. The initial condensation step is carefully controlled to produce a resin that is easily handled during the final fabrication step.

The resole may be produced as a powder, a solution, or a dispersion. It can be compounded with fillers and other ingredients, cast into a mold, coated on a substrate, or used to impregnate a porous material. The thermosetting step then is carried out at temperatures up to 180°C.

If the phenolic reactant is present in excess during the initial condensation reaction, and an acidic catalyst such as sulfuric or oxalic acid is used, then a resin called a novolac is formed.

About 50 percent of the U.S. production of phenolic resin is used for the resole binder used in the manufacture of plywood. The laminates are made by heating the wooden panels coated with the resin under pressure. Phenolics also are used for electric coil varnishes, solvent-based adhesives, and molding powders. Wood fiber, mineral filler, and pigments can be added to the resin prior to the final curing and molding step. Parts such as automobile distributor housings and steering wheels, electrical plugs, and switches are produced from molding powders. Other applications include coatings, friction materials for break linings, and binders for abrasives and foundry molds.

At molecular weights up to 1000, the novolacs are nearly linear; above 1000 branching becomes important. During the subsequent higher-temperature curing step, cross-linking occurs. Unlike the case of resoles, a curing agent that liberates formaldehyde or a readily condensable aldehyde needs to be added in order to convert the novolac to an infusible material. Paraformaldehyde and hexamethylenetetramine ("hexa") commonly are used for this purpose.

Unsaturated Polyester Resins

Unsaturated polyesters are made by condensing polyols with unsaturated dicarboxylic acids or their anhydride derivatives. One or more unsaturated acid components may be

used in the synthesis of the resin, but for economic reasons maleic anhydride is used most often. Fumaric acid produces materials with higher impact resistance, but is more expensive. Aliphatic, aromatic, and hydrogenated aromatic dicarboxylic acid or anhydride comonomers also may be added to the synthesis mixture to modify the product's properties. Phthalic acid often is used because it is inexpensive, improves solubility of the resin in styrene, and increases hardness and resistance to water. Isophthalic acid produces a better-performing product than phthalic acid but is more costly. Aliphatic dicarboxylic acids add flexibility but reduce resistance to water.

Like the dicarboxylic acid portion, a mixture of glycols may be used in resin synthesis. Commonly used are ethylene glycol, bisphenol A, trimethylolpropane, and the various polyethylene and polypropylene glycols. The selection depends on the properties needed in the intermediate resin and in the cured product.

Polymerization is carried out with the neat reactants at temperatures around 200°C and allowed to proceed until a predetermined viscosity or acid value has been attained.

Free radical polymerization inhibitors such as hydroquinone usually are added to the resin mixture to prolong storage life. The presence of these inhibitors causes an induction period in the final curing step. The system will gel after about 1 percent cross-linking takes place, and then the rate of curing rapidly

Styrene-soluble, linear unsaturated polyester

Sixty to eighty percent solutions of the thermoplastic reaction product in monomer solvents such as styrene, vinyl toluene, or methyl methacrylate are packaged and distributed. The resins often are mixed with fillers, initiated with peroxide, and then cured into a hard thermoset material.

increases. Inability to remove the heat generated during this step is a problem in large, thick castings. If the temperature goes too high, the material may become brittle, cracked, and discolored. Up to 20 percent shrinkage can occur during polymerization.

The shrinkage, which causes stress and complicates the casting and lay-up process, can be reduced by using fillers.

The choice of initiator will depend on the fabrication requirements. For low-temperature curing, cobalt–peroxide combinations often are used. The benzoyl peroxide–tertiary amine system is used at low temperatures when oxygen inhibition is a problem. Higher-temperature processes generally rely on the uncatalyzed thermal initiation of initiators such as cumene hydroperoxide or dilauryl peroxide.

Oxygen can severely retard the free radical curing step. This is a major concern in coatings and other applications where the surface-to-volume ratio is high. Cure inhibition produces a tacky surface, or one with low hardness, and chemical and abrasion resistance. In these applications the tertiary amine–benzoyl peroxide accelerator–initiator combination has been found to be useful. Wax also may be added to the formulation. During curing the wax migrates to the surface to form a film that serves as an oxygen barrier. Other ways to minimize oxygen inhibition include thermal post-curing and the use of special resin formulations containing allyl ether or allyl ester functionality. These so-called allyl resins promote post-curing by an autoxidation process.

Flame-resistant products can be made by the inclusion of halogenated glycol or anhydrides in the preparation of the resin. Chlorendic anhydride or tetrabromophthalic anhydride also is used for this purpose. The hydrogen halide that is generated during burning serves as an oxidation inhibitor and promotes charring. The inhibition process is further enhanced by adding antimony oxide or trihalide to the formulation. Because the gases produced by these fire retardants are toxic, a great effort is being made to find nontoxic replacements.

Unsaturated polyesters are fabricated by casting, laminating, pultrusion, and coating. Except for some coating applications, most of these processes incorporate mineral fillers and/or glass fibers to improve strength and lower cost. Treating the glass fiber with an organosilane coupling agent such as 3-methacroylpropyltrimethoxy silane improves adhesion to the polyester and enhances strength and durability even more. Cast products include simulated marble, terrazzo, tile, electrical encapsulants, polymer cements, patching compounds, and bowling balls. Clear unsaturated polyester coatings are used in high-gloss furniture, wood panel, and flooring applications.

Gel coats are a large-volume application for unsaturated polyesters. These are reactive formulations containing pigments, fillers, and thixotropes to prevent sagging of the ungelled coating. The catalyzed resin formulation is sprayed onto a mold and allowed to gel, making an inner layer that forms the surface when the mold is separated. If the mold is glossy smooth, the surface of the coating will be glossy. Several coats may be applied in succession to build the coating thickness and produce special decorative effects. Following application of the surface layer to the mold, the coating is reinforced with a preshaped glass mat using a laminating binder. Alternatively, the reinforcement is applied by spraying a resin formulation containing chopped fibers. The binders used in these applications also might be unsaturated polyester resin formulations. Gel coats are used widely as exterior surfaces for boat hulls, bathroom fixtures, and recreational vehicles.

Filament winding on molds is used to produce large laminated tubes and storage tanks. The continuous glass fiber is guided into a tank filled with resin and then through an orifice to remove excess resin, and finally is wound onto a core. The part is cured in an oven, the core is removed, and additional finishing steps are completed.

Pultrusion is a related process.[37] A continuous, resin-impregnated mat, or a bundle of reinforcing glass, Kevlar, or carbon fibers, is pulled through a heated die where shaping and curing take place. In addition to simple rods, tubing, and sheet, I-beams and complex profiles can be produced. Applications include ladders, reinforcement members for vehicles and buildings, window casements, and moldings of various kinds. There is further

discussion of this process in the section "Fabrication of Plastics" under "Pultrusion."

Large panels are prepared with sheet molding compound (SMC). Typical SMCs contain a polyester resin of high fumaric acid content, calcium carbonate or clay fillers, chopped glass fibers, calcium, magnesium or zinc oxide thickeners, and high-temperature peroxide initiators. Rubbers or thermoplastic resins are used to reduce the amount of shrinkage during curing. SMCs can be compression-molded in minutes at high temperatures and moderate pressures to produce parts with paintable surfaces. Plastic auto and truck body panels are produced this way.[38]

Epoxies

Epoxies are a broad class of reactive monomers and resins containing epoxy (oxirane) functionality. Introduced commercially at the end of World War II, they first were used for coatings, but soon began to be used for adhesives, cast products, and laminating binders. Cured epoxies have good chemical and corrosion resistance, high adhesion to a variety of substrates, and low shrinkage during curing.

The most widely used epoxies are those derived from the condensation of epichlorohydrin with bisphenol A. The degree of oligomerization, can be controlled by varying the epichlorohydrin/bisphenol A ratio. When this ratio is 2:1 or more, the major product is the diglycidyl ether monomer, where n is zero in the general formula.

As the molar ratio of epichlorohydrin to bisphenol A declines, n and the molecular weight of the resin increase. Commercial liquid resins based on bisphenol A and epichlorohydrin have average molecular weights of under 1000. The solid resins are low-melting solids with average molecular weights below 4000, corresponding to an average n of 13. For each epoxy group that reacts, a hydroxyl group is generated. Thus n also corresponds to the number of hydroxyl groups in the molecule. When molecular weights are much higher than 4000, the high viscosity, limited solubility, and sluggish reactivity of the resin restricts its use for cross-linking reactions.

When the molar ratio of bisphenol A and epichlorohydrin approaches unity, resins with molecular weights greater than about 30,000 and n greater than about 40 are produced. These high molecular weight thermoplastics are called phenoxies:[39]

Because phenoxies are not cross-linked and have few unreacted epoxy groups remaining, they can be considered to be bisphenol A–propylene glycol polyethers. They are thermally stable, ductile polymers that can be processed in conventional molding, extruding, and thermoforming equipment. They are made into thermoplastic adhesives, coatings

and sealants, film, and molded articles. The hydroxyl groups in the phenoxies can be esterified and used as reactive sites to prepare cross-linked coatings. The phenoxy esters exhibit excellent properties as paint primers.

With the exception of the phenoxies, most applications for epoxies require the solvent and chemical resistance and heat stability of thermosets. Because unmodified epoxies do not cross-link efficiently at convenient processing temperatures, a curing agent or a hardener is used. Curing agents can be catalytic or coreactive. The most common catalytic agents are Lewis acids such as boron trifluoride or Lewis bases such as tertiary amines. The Lewis acids weaken the carbon–oxygen bond of the epoxy ring by coordinating with the oxygen. This lowers the activation energy for ring opening enough to permit ready cross-linking by hydroxyl groups on other chains. Tertiary amines catalyze ring opening by the following scheme:

Multifunctional primary and secondary aliphatic polyamines are used as coreactants. They add readily to the epoxy group at low temperatures to produce highly cross-linked networks. The aromatic amines are somewhat slower than the aliphatic amines, but provide higher heat stability. Examples of amine curatives include diethylenetriamine, triethylenetetramine, *N*-aminoethylpiperazine, and *m*-phenylenediamine.

Amine curing agents require special precautions. Unreacted amine can migrate to the surface of the product, resulting in odors and possible toxic reactions in sensitized people. To minimize these problems, multifunctional oligomeric amines often are used. Because they have higher molecular weight and functionality, they are less volatile and have a higher probability of being incorporated into the cured product than the lower molecular weight amines.

Mono- and dicyclic acid anhydrides also are widely used as curing agents. Examples include the solids phthalic anhydride and pyromellitic anhydride, and liquid anhydrides such as dodecenylsuccinic anhydride. The reaction rate between an epoxy resin and an anhydride is low, but it can be significantly improved by the addition of a catalytic amount of tertiary amine. Anhydride-promoted curing involves both anhydride and epoxy ring opening steps. The amine preferentially opens the anhydride to form the carboxylate group, which, in turn, reacts with the epoxide group.

One advantage of epoxy thermosets is that the amount of shrinkage that takes place during curing is low compared to that for vinyl addition and condensation polymerization. Low shrinkage means that the polymer has a lower degree of internal strain and retains better contact with a substrate to which it is bonded. Thus, epoxy coatings, adhesives, and encapsulants are noted for good adhesion.

Cured epoxy resins are rigid and brittle. Their impact resistance is improved by using oligomeric amine and polyol coreactants.

Because the equivalent weight per active amine is high in these products, there are fewer cross-link sites per unit. T_g is reduced along with hardness and modulus. Mineral fillers such as mica, quartz, glass, silica, marble powder, and alumina reduce shrinkage, lower thermal expansion, and reduce cost.

Coatings and laminating binders are among the largest uses for epoxies. In electronics they are used for potting and encapsulating and for impregnating cloth and fibrous mats for circuit boards. They also are used to bond carbon filament, high-strength fibers, and other components into aircraft and missile body components, chemical holding tanks, oil pipes, fishing poles, golf club handles, and other products where strength, water resistance, and resiliency are needed.

Silicone Resins

The basic chemical technology of silicones and silicone elastomers is covered in Chapter 18; so the discussion here is limited to silicone resins. These materials have many applications, including use in paints and varnishes, molding compounds, encapsulants, electrical insulation, pressure-sensitive adhesives, laminates, and release coatings. Their heat stability, water repellency, and resistance to solvents and weathering favor their use in these applications.

The production of these resins starts with a mixture that might include di-, tri- and tetrafunctional silanes, such as monomethyl-, dimethyl-, monophenyl-, diphenyl-, monovinyl-, and methylvinylchlorosilane, along with silicon tetrachloride. The tri- and tetrafunctional ingredients are cross-linkers. The resin properties will depend on the chlorosilane mixture, the degree of cure, and processing conditions.

The chlorosilane blend is mixed with inert solvents such as mineral spirits, toluene, or chlorinated hydrocarbons, or esters such as butyl acetate, and then hydrolyzed. Because the hydrolysis reaction can be complicated by the formation of insoluble gels, special care must be taken with respect to the choice and concentration of solvent, temperature, agitation level, sequential addition of the chlorosilanes, water concentration, and use of low molecular weight alcohol modifiers. These variables also affect the molecular structure of the hydrolysate and finished resin properties. The aqueous HCl solution formed by the hydrolysis is separated from the solvent phase, which contains the hydrolysate. The hydrolysate solution then is heated in the presence of catalysts such as metal soaps or acid-treated clays to promote partial condensation or "bodying." Some of the solvent may be stripped off during this step to achieve higher temperatures for more rapid condensation or to promote intermolecular silanol reactions. As the condensation reaction proceeds, the rate drops because of reduced silanol concentration, increased steric hindrance, and reduced molecular mobility. Curing generally is incomplete at the end of the bodying step. The resin then is filtered, curing catalyst is added along with solvent as needed to achieve the desired solids level, and the resulting solution is packaged.[46]

POLYMER SYNTHESIS

Unlike the properties of low molecular weight monomeric substances, the chemical and physical characteristics of high molecular weight polymers are strongly dependent on methods of synthesis. The chain-forming mechanism, polymerization conditions, and type of catalyst or initiator can significantly affect molecular weight, polydispersity, branching, stereochemistry, and end groups in the product. Condensation polymerization and addition polymerization are two major ways that monomers join to form polymers. New polymers also can be made by modifying existing ones.

Condensation polymerization occurs when monomers combine by displacing a small by-product molecule (D) for every new unit added to the growing oligomer:

$$n\,A + n\,\mathrm{B} \rightarrow \text{Polymer} + n\,\mathrm{D}$$

Water often is the by-product of condensation, but other condensation by-products can

include methanol, phenol, and hydrogen chloride, depending on the polymerization reaction. In order for condensation polymerization to occur, the polymerizing molecules must have a functionality of two or more. When the monomers are difunctional, a linear polymer will be formed. If the monomers have a functionality of more than two, cross-linking is possible. Compared to addition polymers, the volume of condensation polymers sold is small. However, they constitute a diverse and important class of materials, which includes polyesters, polycarbonates, polyimides, polysulfides, polysulfones, the phenol-formalde-hyde and melamine resins, and a large percentage of the polyamides and polyurethanes. Examples of most of these materials were discussed in the preceding section.

A characteristic of addition polymerization is that monomers link together by forming new bonds without splitting off by-products. The repeating units have the same atoms as the monomers from which they were derived:

$$\mathrm{M} + n\,\mathrm{M} \rightarrow (\mathrm{M})_{n+1}$$

Both addition and condensation polymerization can take place by chain-growth and step-growth mechanisms. In chain-growth polymerization, chain extension can occur only after an initiation step has taken place. Only the initiating species and the propagating intermediate are capable of reacting with the monomer. One consequence of this is that high molecular weight species are generated early in the reaction, and the monomer still is present until very high conversions to the polymer have occurred.

Chain growth can be initiated by free radicals, acids and bases, and coordination compounds. Because the number of propagating species is never more than the number of initiating species, the average molecular weight of the product will depend in part on the concentration of the initiator. The higher the concentration of the initiating species, the higher the concentration of propagating intermediates, and the smaller the number of monomer units per growing intermediate.

Molecular weight also is determined by the frequency of terminating reactions. In free radical polymerization, the larger the number of initiating and propagating species present at any time, the greater the probability that coupling type terminations will occur. Thus, in polymerization where high molecular weight polymer is needed, small concentrations of initiator are used.

In step-growth polymerization an initiating species is not needed for chain extension, though catalysts may be effective. Monomers and intermediate oligomers have about the same reactivity because they share the same reactive end groups. Monomers react with other monomers and with oligomeric intermediates at nearly the same rate, so monomer concentration declines rapidly. Oligomeric intermediates, in turn, react along themselves and remaining monomer to give higher and higher molecular weight intermediates. However, unlike the case of chain polymerization, very high molecular weight polymer does not form until very late in the reaction.

In step-growth polymerization, the average degree of polymerization (DP), or number of monomer units per molecule, is given by the relationship $DP = 1/(1 - P)$, where P is the fraction of end groups consumed. The following table shows that when half the monomer molecules have reacted, the degree of polymerization is only 2, and it is only 20 when 95 percent have reacted. If a step-growth polymer requires a DP of 100 to have useful properties, 99 percent conversion of end groups is required.

P	DP
0.5	2.0
0.7	3.3
0.9	10
0.95	20
0.99	100
0.999	1000
1.0	∞

Monomer conversions of 99 percent or more often are required for good properties in step-growth polymers; so equilibrium processes must be driven to completion, and polymerization should be nearly free of

competing side reactions. If these conditions are met, polymers with good properties can be made, and molecular weight control can be achieved by adjusting the ratio of the reactants.

Many addition polymerizations are chain-growth processes involving the carbon–carbon double bond, and most condensation reactions are step-growth processes. However, there are notable exceptions. For example, polyurethanes are made by the step-growth addition of glycols and diamines to the carbon–nitrogen double bond of diisocyanates.

Free Radical Addition Polymerization

Free radical chain-growth addition polymerization of vinyl monomers is an important route to commercial polymers. The chain-growth mechanism involves three steps: initiation of a chain, propagation of the growing chain, and termination of the reactive intermediates.

Initiating free radicals can be generated by high-energy particles, oxygen, redox reactions, and thermal- or radiation-induced dissociation of initiators. Thus their sources include UV, electron beam, and autoxidative curing of coatings and adhesives; redox initiation in emulsion polymerization; and thermal decomposition of peroxides and azo compounds in solution, suspension, and bulk manufacturing processes.

Benzoyl peroxide is a frequently used thermal initiator, both for the manufacture of low density polyethylene resin and for unsaturated polyesters. At moderate temperatures benzoyl peroxide disassociates into a pair of benzoyl radicals that, in turn, further decompose into phenyl radicals (I·) and carbon dioxide.

These radicals add, in turn, to the double bond of an ethylene molecule to produce a carbon-centered monomer radical. The double bond becomes a single bond (saturated) and a new radical species is formed.

$$\mathrm{I}\cdot + \mathrm{CH_2{=}CH_2} \rightarrow \mathrm{I{-}CH_2CH_2}\cdot$$

Dissociation and subsequent addition of the initiator fragment to a monomer are together called the initiation step.

The new carbon radical formed during initiation quickly adds to another ethylene molecule. Then, in rapid succession, the chain propagates from the reactive chain end, adding thousands of units in a fraction of a second until all the monomer is consumed, or, what is more likely, a termination or chain transfer reaction intervenes.

$$\mathrm{I{-}CH_2CH_2}\cdot + n\,\mathrm{CH_2{=}CH_2} \rightarrow \mathrm{I{-}CH_2CH_2(CH_2CH_2)}_n\cdot$$

Termination can take place by coupling and by disproportionation reactions. Coupling occurs when two free radical centers bond to each other:

$$\mathrm{{-}CH_2CH_2}\cdot + \cdot\mathrm{CH_2CH_2{-}} \rightarrow \mathrm{{-}CH_2CH_2CH_2CH_2{-}}$$

$$\mathrm{{-}CH_2CH_2}\cdot + \mathrm{I}\cdot \rightarrow \mathrm{{-}CH_2CH_2{-}I}$$

Disproportionation can occur if a radical center of one intermediate abstracts a hydrogen atom from a carbon next to the radical center of another propagating intermediate. A double bond and a saturated carbon are produced at the expense of two reactive intermediates:

$$\mathrm{{-}CH_2CH_2}\cdot + \cdot\mathrm{CH_2CH_2{-}} \rightarrow \mathrm{{-}CH_2CH_3} + \mathrm{{-}CH{=}CH_2}$$

Another way that a chain can stop growing is by chain transfer. During chain transfer, polymerization continues because a new propagating radical is generated. Chain transfer occurs when a propagating free radical abstracts a hydrogen, or some other atom, from a donor. The donor can be a monomer, polymer, initiator, solvent, or another propagating intermediate.

$$-CH_2CH_2\cdot + R-H \rightarrow -CH_2CH_3 + R\cdot$$

$$R\cdot + CH_2{=}CH_2 \rightarrow R-CH_2-CH_2\cdot$$

Chain transfers generally result in branching because hydrogens along a chain's backbone are more likely to be abstracted than hydrogens on a terminal carbon. There are two reasons for this. There are often more internal hydrogens than terminal hydrogens, so the probability of encountering an internal hydrogen is greater. Also, internal hydrogens usually are more weakly bonded than terminal hydrogens. The order of decreasing carbon–hydrogen bond strength is primary > secondary > tertiary > benzyl. Thus, the secondary hydrogens along the backbone of polyethylene are more readily abstracted than the terminal methyl hydrogens.

The highly branched structure of LDPE is a result of both intermolecular and intramolecular chain transfer reactions. Intramolecular chain transfer, called backbiting, takes place when the reactive center bends around and encounters one of its own backbone hydrogens. This often occurs at the fifth and sixth positions from the radical center.

Intermolecular chain transfer

$$\sim CH_2\cdot + \sim CH_2\overset{H}{\overset{|}{C}}HCH_2\sim \longrightarrow \sim CH_3 + \sim CH_2\dot{C}HCH_2\sim$$

$$\sim CH_2\dot{C}HCH_2\sim + CH_2{=}CH_2 \longrightarrow \sim CH_2\overset{CH_2CH_2\cdot}{\overset{|}{C}}HCH_2\sim$$

Intramolecular chain transfer

$$\xrightarrow{CH_2=CH_2} \sim CH_2\overset{CH_2CH_2\cdot}{\overset{|}{C}}HCH_2CH_2CH_2CH_3$$

When chain transfer occurs, the degree of polymerization decreases. In some polymerization processes, substances called chain transfer agents are added to modify the molecular weight and structure of a polymer. When used in large concentrations, they produce low molecular weight polymers called telomers. Thiols are effective chain transfer agents because the hydrogen of the sulfur–hydrogen bond is easily abstracted by a carbon-centered free radical, and the resulting thiol radical readily adds to a carbon–carbon double bond to start a new chain.

$$RS-H + \sim CH_2\underset{\underset{R}{|}}{C}H\cdot \longrightarrow RS\cdot + \sim CH_2\underset{\underset{R}{|}}{C}H_2$$

$$RS\cdot + CH_2{=}\underset{\underset{R}{|}}{C}H \longrightarrow RS-CH_2\underset{\underset{R}{|}}{C}H\cdot \longrightarrow$$

If a hydrogen is easily abstracted from a bond, but, for electronic or steric reasons, the newly generated free radical does not readily add to a monomer to continue the polymerization process, polymerization is inhibited. Substances that form very stable free radicals, such as hindered phenols and amines, fall into this class. Thus, the difference between a chain transfer agent and an inhibitor (I—H) in some instances is a matter of degree.

$$I-H + \sim CH_2\underset{\underset{R}{|}}{C}H\cdot \xrightarrow{fast} \sim CH_2\underset{\underset{R}{|}}{C}H_2 + \underset{stable}{I\cdot}$$

$$I\cdot + CH_2{=}\underset{\underset{R}{|}}{C}H \longrightarrow \text{No, or very slow reaction}$$

Inhibitors often are added to monomers to increase storage life and stabilize the resin during processing or the polymer during normal use. Some do this by scavenging initiating radicals and others by decomposing adventitious peroxides. Commercial hindered phenol and amine antioxidants couple with free radicals generated by autoxidative or thermal decomposition during polymer processing or product aging.

Di-*tert*-Butyl-*p*-cresol

Unreactive

Atmospheric oxygen, which can form peroxide and hydroperoxide initiators, also can serve as a potent polymerization inhibitor; so it is rigorously excluded from most free radical polymerizations. The rate of coupling between molecular oxygen and a free radical intermediate during polymerization is several orders of magnitude greater than the rate of propagation of the new oxygenated intermediate. Thus oxygen reduces the polymerization rate, lowers average molecular weight, and incorporates undesirable peroxide links in the product.

When an unsymmetrical monomer such as styrene or vinyl chloride is polymerized, two propagating species are possible. A combination of steric and electronic factors will favor one of the two possibilities and will determine the stereochemical course of the reaction (see above, section on "Polymer Structure and Nomenclature"). This results in "head-to-tail" linking. In coupling-termination reactions, however, head-to-head linking occurs.

$$\sim CH_2\underset{\underset{R}{|}}{CH}\cdot + \cdot\underset{\underset{R}{|}}{CH}CH_2\sim \longrightarrow \sim CH_2\underset{\underset{R}{|}}{CH}—\underset{\underset{R}{|}}{CH}CH_2—$$

Copolymerization is similar to homopolymerization, but the reactivity of the different monomers toward copolymerization and competing homopolymerization may differ considerably. For a system with two different monomers the following reactions are possible:

$$M_1\cdot + M_1 \xrightarrow{k_{11}} M_1—M_1\cdot$$
$$M_1\cdot + M_2 \xrightarrow{k_{12}} M_1—M_2\cdot$$
$$M_2\cdot + M_2 \xrightarrow{k_{22}} M_2—M_2\cdot$$
$$M_2\cdot + M_1 \xrightarrow{k_{21}} M_2—M_1\cdot$$

The first and third reactions are homopolymerizations, whereas the second and fourth are copolymerizations. The rates of these competing processes are given by the rate constants, k. A term called the reactivity ratio, r, where $r_1 = k_{11}/k_{12}$ and $r_2 = k_{22}/k_{21}$, is a measure of the relative tendency of

the monomers to homopolymerize or copolymerize. If $r_1 > 1$, then M_1 tends to homopolymerize; whereas if $r_1 < 1$, copolymerization is preferred. If r_1 and r_2 are both equal to 1, there is no preference, and the resulting polymer will have a random distribution of monomer units. When r_1 and r_2 are both near zero, as is the case with maleic anhydride and styrene, an alternating copolymer is formed. There is a tendency toward alternation when the product of r_1r_2 approaches 0, and a tendency to produce random copolymers when r_1r_2 approaches one. When r_1 is much greater than one and r_2 is much less than one, as with styrene and vinyl acetate ($r_1 = 55$ and $r_2 = 0.01$), a nearly pure homopolymer of M_1 is produced. Table 19.4 is a partial list of reactivity ratios.[40] Reactivity ratios depend on the reaction temperature and on the reaction type; and they will also be different for free radical, anionic, and cationic copolymerizations.

Usually reactivity ratios fall between 0 and 1, so there is some tendency to alternate. Because the monomers in a copolymerizing batch reaction are seldom consumed at the same rate, the monomer composition of the reacting mixture will change rapidly. Thus, it is often necessary to adjust for compositional drift by adding monomers continuously to the reaction mixture.

TABLE 19.4 Some Free Radical Copolymerization Reactivity Ratios

	r_1	r_2
Ethylene		
1-Butene	3.60	0.16
Butyl acrylate	0.01	13.94
Propylene	3.20	0.62
Vinyl chloride	0.21	0.01
Methyl methacrylate		
Acrylamide	3.00	0.90
Isoprene	0.26	0.65
Maleic anhydride	5.20	0.02
Styrene	0.48	0.58
Vinyl acetate		
Methyl acrylate	0.03	6.38
Acrylonitrile	0.06	5.51
Ethylene	1.03	0.88
Vinyl chloride	0.25	1.64
Vinyl chloride		
Methyl acrylate	0.00	5.00
Isobutylene	2.12	0.27
Styrene	0.16	23.92
Vinylidene chloride	0.20	1.80

The relationship of the molar ratio of monomers in a copolymer, n, and in the monomer mixture from which the copolymer is formed at any instant, x, is given by:

$$n = [r_1x + 1]/[(r_2/x) + 1]$$

Depending on the values of r_1 and r_2, there is a monomer composition that will yield a copolymer of the same composition. This composition, known as the azeotrope, can be derived from the above equation where $n = x$:

$$n_{az} = [M_1/M_2]_{az} = (1 - r_2)/(1 - r_1)$$

In principle, a pair of monomers at the azeotropic composition will produce a copolymer of the same composition throughout the batch reaction.

Copolymerization is an important way to produce properties that are not possible with homopolymers. For example, the homopolymer of vinylidene chloride is highly crystalline, and though it has excellent moisture and oxygen barrier properties, it does not produces very strong film or fiber. Copolymerization with 15 percent vinyl chloride disrupts the regular structure of the homopolymer to produce a stronger, clearer, more flexible material. The copolymer retains much of the barrier properties of the homopolymer and finds wide use for food packaging and filament. Other commercial copolymers include styrene–acrylonitrile, discussed above; vinylidene fluoride–hexafluoropropylene, a heat- and oil-resistant elastomer; styrene–butadiene rubber; ethylene–vinyl acetate hot melt adhesive; and 2-ethylhexyl acrylate–vinyl acetate–acrylic acid pressure-sensitive adhesives.

Ionic Chain Addition Polymerization

Some monomers can polymerize by ionic processes. Table 19.5 lists types of chain extension that some monomers undergo to produce homopolymer. Ionic polymerizations

TABLE 19.5 Type of Initiation Suitable for a Select Group of Monomers toward Homopolymerization

Monomer	*Anionic*	*Cationic*	*Free Radical*	*Transition Metal Catalysis*
Acrylamide	Yes	No	Yes	No
Acrylic esters	Yes	No	Yes	No
Acrylonitrile	Yes	No	Yes	No
Butadiene	Yes	Yes	Yes	Yes
Caprolactam	Yes	Yes	No	No
Diallyl phthalate	No	No	Yes	—
Epoxies	Yes	Yes	No	No
Ethylene	No	No	Yes	Yes
Formaldehyde	Yes	Yes	No	No
Isobutene	No	Yes	Yes	Yes
Styrene	Yes	Yes	Yes	Yes
Tetrahydrofuran	No	Yes	No	No
Vinyl chloride	No	No	Yes	No
Vinyl esters	No	No	Yes	—
Vinyl ethers	No	Yes	No	—
Vinylidene chloride	No	No	Yes	—
Vinylidene fluoride	No	No	Yes	—

of vinyl monomers are similar in some respects to free radical chain polymerizations. One difference is that counterions are present during ionic polymerization. Depending on whether the counterion is weakly or strongly associated with the growing chain, it can markedly influence the rate of polymerization and the stereochemical result. Another difference is that the polarity of the solvent plays a larger role in ionic additions than it does in nonionic processes.

Cationic Polymerization. In cationic chain polymerization, the propagating center is positively charged. Protonic acids such as HCl and H_2SO_4 or Lewis acids such as BF_3, $TiCl_4$, I_2, and $AlCl_3$ have been used. Lewis acid catalysts usually require trace amounts of a cocatalyst such as water to be effective. Cationic polymerization also can be initiated photochemically by using onium salt photoinitiators. These substances generate protons during irradiation with UV light and are used in coating and adhesive applications.

When initiation takes place with an unsymmetrical monomer such as 2-methylpropene, addition occurs to the double bond to give the more thermodynamically stable carbonium ion. Propagation continues in like manner, always producing the more stable intermediate. As in free radical polymerization, this results in head-to-tail type chain growth.

Reaction rates are determined by both electronic and steric factors; but unlike the case of free radical polymerization, rates are strongly dependent on the dielectric characteristics of the solvent and basicity of the counterion. Rates of addition to simple monomers are usually in the order of increasing stability of the carbonium ion intermediate. Methyl groups or substituents that can delocalize a positive charge undergo

polymerization more rapidly; thus the following order of decreasing rates of vinyl polymerization is expected:

$$CH_2{=}CHOR > CH_2{=}CHPh > CH_2{=}C(CH_3)_2 > CH_2{=}CHCH_3 > CH_2{=}CH_2$$

Like free radical polymerization, cationic polymerizations undergo chain transfer and termination reactions. Chain transfers can take place by proton transfer to monomer, by hydride transfer from a chain, and by a variety of alkylations and cyclizations. Termination by coupling two cationic centers is not possible, but termination can occur when propagating centers react with water or other basic contaminants to give stable products.

Some rubbers, acetal resins, and a variety of polyethers are prepared on a commercial scale by cationic polymerization. Vinyl ethers are among the easiest monomers to polymerize by the cationic process. The carbonium ion is readily formed because the pendant oxygen participates in delocalization of the positive charge. When polymerization of isobutyl vinyl ether is carried out at temperatures in the range of $-80°C$ using boron trifluoride etherate as initiator, isotactic crystalline polymer can be produced. Ethyl and isobutyl ethers produce polyvinyl ethers with low T_g's, and are used in pressure-sensitive adhesives.

Butyl rubber is made by low-temperature cationic copolymerization of isobutylene with 1 to 10 percent isoprene. The double bonds left in the polymer are cross-linking sites for subsequent vulcanization.

Because of its low permeability to gases, butyl rubber is used for inner tubes, gaskets, tubing, and sealants.

Other cationic processes include ring opening polymerization of epoxies, discussed below, and acid-catalyzed polymerization of aldehydes, discussed above.

Anionic Polymerization. Some monomers undergo chain extension when initiated by bases. As in cationic processes, both vinyl addition and ring opening chain extension can occur. The propagating species in these reactions are negatively charged. Monomer substituents attached to the carbon–carbon double bond that stabilize a negative charge, such as cyano, nitro, chloro, phenyl, and carbonyl groups, readily undergo anionic polymerization. Examples include acrylonitrile, methyl methacrylate, styrene, and butadiene, in order of decreasing reactivity. Rates are dependent on the basicity of the initiator, the ability of the solvent to solvate the reactants, and the counterion.

Alkali metals initiate anionic polymerization indirectly by an electron transfer process:

$$Na\cdot + CH_2{=}CHR \longrightarrow Na^+ + \cdot CH_2\bar{C}HR$$

The anion radical generated in the transfer step rapidly dimerizes to produce a dianion that serves as a difunctional initiator:

$$R\bar{C}HCH\cdot + \cdot CH_2\bar{C}HR \longrightarrow {}^{-}\underset{\underset{R}{|}}{C}HCH_2{-}CH_2\underset{\underset{R}{|}}{C}H^{-}$$

Strong bases such as imide, alkoxide, cyanide, and alkylate ions (e.g., butyl lithium) that add directly to the monomer's double bond constitute a second class of anionic initiators:

$$NH_2^- + CH_2{=}CHR \longrightarrow NH_2CH_2\underset{\underset{R}{|}}{C}H^-$$

For polymerization by amide ion, NH_2^-, the following order of rates of polymerization occurs:[41]

$$CH_2{=}CHCN > CH_2{=}CCO_2CH_3 > CH_2{=}CHC_6H_5 > CH_2{=}CHCH{=}CH_2$$

Like other chain-addition processes, anionic chain polymerization is complicated by competing chain transfer and termination steps. However, there are conditions where termina-

tions are unfavorable. If there is rigorous exclusion of oxygen, water, carbon dioxide, and other substances that can quench the reactive intermediate, a "living polymer" is possible. When this occurs, all of the monomer has been consumed, and only the potentially reactive macroanions remain. If additional monomer is added at this point, propagation will resume.

The production of block copolymers involves synthesis of a living homopolymer, followed by addition of a second monomer. When the second addition is complete, a living diblock macroanion is formed. A third monomer can be added to the sequence, giving a triblock macroanion.

$$I^- + xM \longrightarrow I\text{-}(M)_x^- \xrightarrow{yA} I\text{-}(M)_x\text{-}(A)_y^-$$

Homo-macro anion Diblock macro anion

$$\xrightarrow{zN} I\text{-}(M)_x\text{-}(A)_y\text{-}(N)_z^- \xrightarrow{H^+}$$

Triblock macro anion

$$I\text{-}(M)_x\text{-}(A)_y\text{-}(N)_z\text{-}H$$

Triblock copolymer

To generate the unreactive polymer, the macroanion intermediates can be quenched with a proton donor. Styrene–butadiene–styrene and other triblock copolymers are commercially available.

Ring Opening Addition Polymerization

Ring opening polymerizations also can occur by chain-growth and step-growth processes. Depending on the specific monomer, ring opening additions can be initiated by acids, bases, coordination catalysts, and free radicals. Because cyclic monomers can be highly strained, the driving force for the reaction is attainment of the more thermodynamically stable open-chain product. In addition to the epoxies already discussed above, other monomer classes that undergo ring opening addition polymerization include: tetrahydrofurans, trioxanes, episulfides, alkyleneimines, lactones, lactams, and cyclic siloxanes.

Commercial polyethers are made from ethylene oxide and propylene oxide by high-pressure, acid- or base-catalyzed processes. Initiation by base occurs by ring opening addition to the epoxy group and subsequent attack by the alkoxide ion generated on another monomer molecule. If sodium hydroxide is the initiator, the product is a diol. If an alkoxide is the initiator, one of the chain ends is an ether group, and the other is a hydroxyl group.

When a Bronsted acid initiator such as sulfuric acid is used, a low molecular weight polyethylene oxide diol is produced. Protonation of the cyclic ether oxygen is followed by ring opening displacement on the monomer.

The lower molecular weight polyethylene oxides are viscous, water-soluble liquids or waxy solids used as surfactants, lubricants, humectants, and cosmetic and pharmaceutical formula bases. Metal oxides and organometallic initiators, such as aluminum and zinc alkyls, produce higher molecular weight polyethylene oxides. The resins are low-melting, easy-to-process solids that are soluble in a broad range of solvents. They find use as thickeners, water-soluble packaging, and latex stabilizers.

Propylene oxide undergoes similar ring-opening polymerization. The products are less soluble in water than the corresponding polyethylene oxides. They serve as toughening soft segments in polyurethanes when they are polymerized with oligomeric diisocyanates.

Alkylenimines take part in acid-catalyzed ring opening polymerization, like the epoxies. Coordination of a proton or Lewis acid with the ring nitrogen weakens the already strained carbon–nitrogen bond. Propagation occurs when the nucleophilic nitrogen of a monomer collides with a coordinated monomer to open up the ring.

Ethylenimene

A type of ring opening addition polymerization using transition metal catalysts not unlike Ziegler-Natta catalysts is called metathesis polymerization.[42] Metathesis polymers, first made in the 1960s, are called polyalkenamers. Polymers made from cyclooctadiene and norbornadiene have applications as specialty elastomers. The polymerization of norbornadiene can be summarized as follows:

Norbornadiene

Polymer Modification

In principle, polymers can undergo the same reactions that simpler molecules undergo. Much use has been made of this capability to produce polymers that cannot be made directly from monomers. Substitutions and hydrolysis of pendant groups produce new functional groups, and controlled degradation reduces molecular weight. In addition, physical processes that change molecular orientation and morphology also can be thought of as polymer modifications. Some examples were provided in discussions about grafting, cross-linking, and film orientation; others are given here.

The conversion of cellulose into useful products is an illustration of how degradation, functionalization, and molecular orientation of a polymer can be applied to commercial advantage. Cellulose is the most common naturally occurring organic polymer. Over a trillion pounds of it are converted annually into paper and other cellulose-based products.

Cellulose

Although it is linear, and nominally thermoplastic, cellulose is not readily soluble and will not flow below its decomposition temperature.

The strong intermolecular hydrogen bonding in the polymer can be reduced by replacing one or more of the hydroxyl groups in the repeating unit by ester, ether, nitro, or xanthate groups.

where R = —NO_2 Nitrate

= —$\overset{O}{\overset{\|}{C}}CH_3$ Acetate

= —CH_3 Methyl ether

= —$\overset{S}{\overset{\|}{C}}$—$S^-Na^+$ Xanthate, sodium salt

The degree of substitution can vary from 0 to 3.

Several important products are made from cellulose this way. The insoluble polymer is dissolved by partially degrading it in alkali solution. In the viscose process, carbon disulfide is added to the basic aqueous mixture to convert some of the hydroxyl groups to xanthate groups. The solution produced this way is allowed to age until it thickens, and it then is extruded as filament into a bath that regenerates the free hydroxyl functionality. During processing the filament is stretched to orient the polymer and improve its strength. The yarn, called rayon, is used for textiles and tire cord. Cellophane is made when the viscose is cast onto a roll or extruded through a narrow slit into an acid bath, and then plasticized with glycerine.

The partially esterified acetate, propionate, and butyrate esters of cellulose also have proved to be useful. They can be made with a wide range of properties because of the latitude provided by varying molecular weight, ester content, and amount and type of plasticizer. The cellulose esters have been made into fiber, thickeners for coatings, and molded articles such as toys, combs, eyeglass frames, and cutlery handles. They have been extruded into pipe, formed into sheet, and used for a variety of decorative and protective laminates.

The methyl and ethyl ethers of cellulose have properties that are similar to those of the ester derivatives, and they have been used for related applications. Ethyl cellulose is injected-molded into parts with good low-temperature toughness and used as an additive in coatings and adhesives to improve flexibility. The methyl derivative is used as a thickener and emulsifier for paints, adhesives, food, and pharmaceutical products.

The conversion of polyvinyl acetate into a family of polyvinyl alcohol (PVOH) resins is an important modification of a synthetic polymer. PVOH cannot be made directly from vinyl alcohol because the stable form of vinyl alcohol is its tautomer, acetaldehyde:

$$CH_2{=}CHOH \rightarrow CH_3CH{=}O$$

Thus, polyvinyl alcohol is made by basic hydrolysis of polyvinyl acetate.

The solubility of PVOH in water depends in a unique way on its degree of hydrolysis. Maximum solubility occurs at around 88 percent hydrolysis. At higher levels, the amount of intermolecular hydrogen bonding becomes high enough to reduce solubility. PVOH has a good combination of tensile strength and solvent resistance and has been spun into fibers, made into film, and used for coatings and adhesives.

Reaction of PVOH with butyraldehyde or formaldehyde gives products of commercial significance. Two of these are polyvinyl butyral (PVB) and polyvinyl formal. These are examples of useful resins made by modifying modified polymers. The latter is made into a varnish for coating wire and cable, and the former is extruded into sheet. Large quantities of PVB are used for the energy-absorbing layer in auto windshields.

Polyvinyl butyral

Polyvinyl formal

There are many examples where a material's bulk properties are adequate for an application but its surface properties are not. Inks, coatings, or adhesives that do not adequately wet the surface of a substrate might cause printing problems and poor adhesion. Sometimes wetting and adhesion can be improved by the use of primers; but in some cases primers are not helpful, and it is necessary to modify the material's surface chemistry. If the polymer has hydroxyl functionality, coupling agents may be useful. Coupling agents such as 3-aminopropyltrimethoxysilane bind to the surface hydroxyls by forming silylether linkages. The other end of the coupling agent extends from the surface to bond with the coating.

In applications where coupling agents are not effective, other methods can be used to modify surfaces. Nonpolar surfaces can be converted to polar surfaces by chemical, radiation, and plasma-induced oxidations, or by plasma deposition of thin adhering coatings. Modifications such as this increase fabrication costs, but they produce uniquely suitable, high-value products.

Polymerization Methods

Bulk Polymerization. Bulk or mass polymerization is conducted with undiluted monomers and prepolymers. It is most successful in reactions where the monomer, like styrene and methyl methacrylate, will dissolve its polymer. The major difficulty associated with bulk polymerization is dissipating the exotherm that accompanies the process. For example, conversion of a carbon–carbon double bond to a single bond evolves about 30 kilocalories/mole of double bond. For polystyrene this amounts to about 290 calories/g, enough energy to severely overheat the mass if a way is not provided to transfer this heat from the polymerizing bulk.

Heat transfer is relatively simple to manage in small reaction vessels or in processes where the surface-to-volume ratio is high. Bulk polymerizations usually are successful for casting small objects, for making cast sheet, film, and coatings, or in reaction-injection molding. Large-scale bulk polymerization, however, must be carried out with specially designed equipment and processing methods. Heat removal generally can be done by a circulating coolant in a reactor jacket, cooling coils immersed in the reactants, and/or recirculating the reactant mixture through an external heat exchanger. If one or more of the monomers is sufficiently volatile (e.g., acrylonitrile), cooling by refluxing may be feasible if conversion to polymer does not get too high. This method can be very efficient. Adding cold feed to the reacting mass is often an effective means of controlling temperature.

The design and the operation of a bulk polymerization reactor and its agitation system go beyond considerations of temperature control and viscosity, however. The resultant mixing behavior also can affect molecular weight, reaction rate, and copolymer composition, for example. With continuous reactors the degree of backmixing is

a very important consideration. In a completely backmixed reactor, for example, the contents are so thoroughly blended that their composition and conversion are uniform throughout. Such reactors often are employed to achieve a uniform copolymer composition, especially if that composition is nonazeotropic. However, a reactor of this type is not practical for operation at very high conversions, where the rate of polymerization is slow, as the reactor size becomes excessive. At the other end of the backmixing scale is the plug flow reactor, where conversion tends to increase continuously as the reactants pass through. Tubular reactors often permit a close approach to plug flow conditions. Plug flow reactors will be employed where a high-conversion effluent is required. Continuous processes may employ as many as three or more reactors in series, each designed to satisfy particular requirements as the reactant mass progresses to higher levels of viscosity and conversion. Multiple reactors with different design features often are encountered for analogous purposes with batch polymerizations.

Viscosity generally has little effect on the overall chain reaction rate until it reaches a value where an autoacceleration called the Trommsdorff or gel effect starts. The conversion at which this occurs, and its severity, will vary, depending on the polymerizing monomers, temperature, and other reaction variables. The gel effect takes place when the thermal motion of the bulky macromolecules has become so restricted that competing coupling termination reactions are hindered. However, because the monomers still can diffuse readily to the growing macromolecules, the reaction rate will accelerate. Dissipating the resulting heat surge can be difficult, especially in large reactors.

At higher conversions, the concentration of the diffusing monomer species starts dropping off. The reaction rate thus will tend to slow down, and heat generation diminishes. However, the higher viscosity that has developed increases the difficulty of agitation, so heat transfer problems still may persist. Agitators for high-viscosity reactants are slow-moving, to minimize power consumption (which can add significantly to the heat-removal load), and are intended primarily to achieve some blending and to avoid development of stagnant regions where reactants might degrade.

Reaction rates at high conversions ultimately will slow down so drastically that it generally is impractical to complete the polymerization in the reactor. Although post-curing may be carried out at some point downstream in the manufacturing process, unreacted monomers most often are removed by devolatilization. The devolatilized exiting polymer melt then is directly extruded to form pellets for subsequent fabrication steps. In some instances the melt may be fabricated directly into the final product.

Solution Polymerization. In solution polymerization, a solvent is used to maintain a homogeneous mixture of monomers, initiators, catalysts, and polymers throughout the reaction. Compared to bulk polymerization, the lower viscosities with this method generally allow improved heat and mass transfer, as well as faster reaction rates at high conversions.

There are a number of reasons why other methods might be preferred for the production of industrial resins. The use of solvents can lower initial reaction rates and reduce molecular weight to undesirable levels due to chain transfer reactions. In addition, solvents often are expensive, flammable, and toxic, and efficient means must be devised to recycle or dispose of them. Devolatilization with large-scale solution processes can be costly because of high capital and energy requirements.

There are, however, important solution polymerization processes. Examples include processes for making poly(vinyl acetate) and copolymers of vinyl chloride and of styrene. Another is the manufacture of unsaturated polyester resins, discussed above. In this case, the solvent for the resin, usually styrene, is a reactive diluent for subsequent polymerization. The styrene copolymerizes with the resin during the curing process, making it unnecessary to devolatilize the polymer or recycle solvent.

Emulsion Polymerization. Water, being nonflammable, nontoxic, and moderately volatile and having low viscosity, can be an excellent medium for carrying out polymerizations. Because most organic monomers are insoluble in water, suitable dispersal means are needed to utilize an aqueous medium. One way this can be done is by emulsifying the monomer and carrying out polymerization in the emulsified particles. This is done primarily via batch processes. The second method, suspension polymerization, is discussed in the next section.

When emulsifiers, such as soaps, alkyl sulfates, and alkyl sulfonates, are present in water above a minimum concentration called the critical micelle concentration, the emulsifier molecules aggregate into micelles. In micelles, the hydrophobic ends of the emulsifier are clustered at its center, and the hydrophilic ends extend into the aqueous phase. Micelles are spherical or lamellar with dimensions on the order of 2.5 to 5 nm, depending on the emulsifier used.[43] There may be 10^{17} to 10^{18} micelles/cm^3 of water. When monomer initially is added, a small fraction of the molecules is absorbed into the micelles, swelling them slightly. Most of the monomer is dispersed as droplets 1000 to 3000 nm in diameter, stabilized with emulsifier molecules along their surface. They number about 10^{10} to 10^{11} particles/cm^3 of water. These droplets serve as the monomer reservoir for the polymerization. A representation of emulsion polymerization early in the process is shown in Fig. 19.14.

Once the monomer phase is dispersed into droplets, the rate of agitation does not ordinarily affect particle size or particle size distribution. The initiator, which is water-soluble, decomposes thermally in the aqueous phase, producing free radicals that diffuse into the micelles and initiate polymerization there. Termination reactions can occur in a number of ways, but because there are very few growing chains within a micelle, the probability for termination by coupling is low; thus the molecular weight of the polymer can become very high. Moreover, the emulsion has low viscosity throughout the polymerization, so

Fig. 19.14. Schematic representation of emulsion polymerization: M = solubilized monomer; R = initiator, monomer or low molecular weight polymer radical.

efficient heat removal can be achieved, permitting high polymerization rates. Initiators, such as potassium persulfate, are effective at moderate temperatures, generally 70°C or less. Thus, the technique of emulsion polymerization offers a number of advantages. Also, the molecular weight can be lowered as needed with chain-transfer agents. The final product, called a latex, is a stable dispersion of solid polymer particles.

One of the limitations inherent in emulsion systems is difficulty in separating the polymer from the emulsifiers used in synthesis. Where the polymer is isolated from the latex, purification is done by adding an acid, salt, or some other substance that reduces the stability of the emulsion.[44] Except for washing the coagulated latex with water, polymers made in this way—such as some grades of styrene–butadiene copolymer and poly(vinyl chloride)—are not purified extensively. In many respects this lack of purification restricts their use because residual emulsifier in the polymer can impair the polymer's electrical properties, increase moisture sensitivity, and reduce clarity. Another difficulty is environmental: unreacted monomers and solvents used in the purification step need to be properly handled, and large quantities of water left from the coagulation step must be purified to meet local waste water regulations.

However, many products are formulated and used without isolating the polymer from the latex. These products include latex paints

and adhesives, textile finishes, paper binders, and some waxes and polishes.

Suspension Polymerization. Suspension and emulsion polymerization are alike in that they are carried out in an aqueous medium. However, in terms of the reaction chemistry and kinetics, suspension polymerization has far more in common with bulk polymerization.[45] A principal advantage of the suspension process is that the heat of reaction can be very effectively transferred from the small polymerizing monomer droplets (generally 50–200 μm in diameter) to the surrounding water, which is in turbulent agitation.

As with bulk polymerizations, monomer-soluble initiators and catalysts are used, but a suspending agent must be added to prevent the droplets from coalescing during polymerization. Unlike emulsifying agents, which markedly lower interfacial tension, many suspending agents are high molecular weight, water-soluble polymers that do not affect droplet surface tension significantly. These agents include gelatin, water-soluble gums, polyvinyl pyrrolidone, polyvinyl alcohol, sulfonated polystyrene, and acrylic acid copolymers. Hydrophobic groups in the polymer locate themselves at the monomer/water interface, increasing viscosity at the surface of the monomer droplet. This effect inhibits coalescence. Another class of suspending agents is based on certain finely divided solids such as calcium phosphate, barium sulfate, bentonite clay, and aluminum hydroxide, which are believed to form a coating on the droplet surface that prevents coalescence.

Agitation intensity is important because it affects both droplet size and size distribution. In contrast to emulsion polymerization, the beads produced by the suspension process can easily be filtered or centrifuged from the final reaction mixture.

A simplified flow sheet of a polystyrene suspension polymerization process is shown in Fig. 19.15. Commercial-scale reactors are generally of 4000-gallon capacity or larger. They frequently are glass-lined although larger reactors are now commonly stainless-steel-clad. The reactors are jacketed for water cooling and equipped with turbine agitators and baffles to obtain good vertical movement and uniform turbulence in the suspension. Reflux cooling is possible with certain systems.

The monomer, which should have very low water solubility, is metered to the reactor, to which deaerated and deionized water has been charged. The ratio of water to monomer is generally in the range of 0.9 to 2:1. Catalyst,

Fig. 19.15. Schematic of a styrene suspension polymerization process. (*Reprinted by permission from* Chem. Eng., **65**, *100 (Dec. 1958); by McGraw-Hill, Inc.*)

chain transfer agent, tints, suspending agent, stabilizers, plasticizers, and other additives usually are charged at this time although delayed addition of one or more of these additives is not uncommon. The reaction is almost always run in a nitrogen atmosphere. The batch then is heated through a predetermined temperature cycle, with the nitrogen pressure maintained to prevent uncontrolled boiling.

When the elevated portion of the temperature cycle is complete, unreacted monomer may be removed from the batch by distillation. The batch then is cooled below the glass transition temperature of the beads and discharged to a hold tank, from which it passes to a continuous centrifuge that removes most of the suspension liquor and washes the beads with water, primarily to remove suspending agent residues. The beads then pass to a dryer, typically a cocurrent, hot air, rotary type.

Suspension processes have certain disadvantages. Except for special applications such as expandable polystyrene and ion exchange resins, the product beads are not favored for critical fabrication steps because their rolling behavior often causes feeding difficulties with extruders and injection-molding machines. To minimize these problems, the beads often are blended with other ingredients or extruded into pellets before being fed to these machines. In addition, there are low levels of suspending agent residues normally found in the product, which often affect optical clarity. More importantly, costs are continually escalating for treating the large volumes of contaminated water discharged from these processes, as well as for the control of organic vapor emissions. The trend, then, is to replace them wherever possible with bulk processes that can avoid these disadvantages.

Gas Phase Fluid Bed Processes. These processes are of particular commercial importance with polyolefin production. A prominent example of this technology is the Unipol process, first commercialized in 1975.[47] A flow sheet is shown in Fig. 19.16. A gaseous monomer, such as ethylene or propylene, along with comonomers and chain transfer

Fig. 19.16. Unipol process for polyolefin manufacture. (*Courtesy Union Carbide.*)

agent, is fed continuously into the bottom of a reactor containing polymer granules maintained in a fluidized state by the gas stream. A transition metal catalyst is added separately. Reactor conditions are relatively mild: 50 to 100°C and 100 to 600 psi. Because of the large ratio of surface to volume of the granules where the polymerization takes place, as well as the turbulent conditions in the fluidized bed, there is very efficient transfer of the heat of reaction to the gas stream. As overheating of the granules can be so effectively avoided, there is very little tendency for them to agglomerate, and product quality control is greatly enhanced.

After passing through the fluidized bed, the gas enters an expanded section of the reactor where fine particles are disengaged. On exiting the reactor, the gas is pumped through a cooling heat exchanger and then is recycled to the reactor. The granular product is removed continuously from the reactor through a gas lock chamber. The small amount of residual monomer in the granules at this point is purged off for recycling. The devolatilized product granules, averaging 0.02 to 0.04 inch, then are conveyed from the system. With polyethylene, monomer yields are in the 97 to 99 percent range. Because the catalysts have very high productivity (10^5–10^6 lb polymer/lb transition metal), catalyst residue levels in the product are so low that it is unnecessary to remove them.

With polypropylene, catalyst selectivity for the isotactic isomer is so high that removal of the small quantity of atactic polymer formed is not needed.

This process has a high level of environmental acceptability in view of its very low emissions of hydrocarbons and other waste streams.

POLYMER RHEOLOGY

Polymer melts and concentrated polymer solutions have many flow properties that distinguish them from monomeric liquids and solutions, and these properties can figure very importantly in the design and operation of process and fabrication equipment.

A principal difference involves the effect of shear on viscosity. Monomeric liquids and solutions exhibit newtonian behavior; that is, viscosity at constant temperature is unaffected by shear rate. Most polymer melts and concentrated polymer solutions, however, are nonnewtonian; their viscosity at a given temperature will change as the shear rate is increased. In general, it will drop as illustrated in the viscosity versus shear rate plot in Fig. 19.17. At very low shear rates there is a newtonian plateau where viscosity does not change. Depending on molecular weight and structure, a shear rate is reached where viscosity starts to drop off. This dropoff continues until a second newtonian plateau is reached. The second plateau, however, has no practical significance because of the extremely high shear rates needed to reach it and the virtually unavoidable thermal effects that tend to mask it. The nonnewtonian portion of the curve, on the other hand, is so frequently encountered in operations with polymers and concentrated polymer solutions that it cannot be overlooked by designers and operators.

Fig. 19.17. Typical log viscosity vs. log shear rate plot for shear-thinning fluids.

Because the nonnewtonian region in Fig. 19.17 usually can be reasonably approximated by a straight line on a log–log plot, the "power law" equation is a particularly straightforward and useful tool for modeling this behavior. It can be expressed by:

$$\eta = \eta_0(\gamma/\gamma_0)^{(n-1)} \qquad (19.1)$$

where η = apparent viscosity, γ = shear rate, n = the flow behavior index, and subscript zero denotes reference values, often taken at the onset of nonnewtonian behavior. It can be seen that where $n = 1$, there is no shear rate effect, and behavior therefore is newtonian. For most polymers and concentrated polymer solutions, n is in the range of 0.25 to 0.5. Where $n > 1$, the viscosity increases with shear rate, but such shear-thickening materials are not encountered in practice with polymer solutions and melts.

Many fluids also exhibit time-dependent viscous effects. With some pastes, drilling muds, and latex paints, for example, there is a drop in viscosity with time when they are sheared at a constant shear rate. This effect, known as thixotropy,[48,49] is illustrated in Fig. 19.18. There is an opposite effect, rheopexy, where there is an increase in viscosity with time under shear. Fluids with this property are relatively rare, however.

Viscosity generally decreases markedly with temperature. An Arrhenius-type equation commonly is used to relate viscosity with temperature in ranges well above T_g:

$$\eta = Ae^{B/T} \qquad (19.2)$$

where A and B are constants characteristic of the material, and T is the absolute temperature. Estimation of the constants requires viscosity data at two temperatures. For

Fig. 19.18. Viscosity vs. time relationship for a thixotropic fluid under constant shear until $t = t_f$.

temperatures in the range

$$T_g < T < (T_g + 100°C),$$

the Williams-Landel-Ferry equation[50] is more reliable:

$$\log(\eta/\eta_{T_g}) = -17.4(T - T_g)/[51.6 + (T - T_g)] \quad (19.3)$$

The viscosity of polymer melts at zero shear is a function of the weight average molecular weight, M_w, where the molecular weight distribution (MWD) is narrow:

$$\eta_0 = KM_w^j \quad (19.4)$$

For concentrated polymer solutions, the following related equation is used:

$$\eta_0 = K'(cM_w)^j \quad (19.4a)$$

where η_0 = zero-shear viscosity, K and K' are proportionality constants, and c = polymer solution concentration. The exponent j is unity for low M_w, and 3.4 where M_w exceeds a critical molecular weight, M_c, at which there is a shift to a more restricted mode of chain movement. M_c is a function of the solvent used (if any) and the composition and structure of the polymer; it usually is in the range of 5000 to 20,000.[51] As the MWD broadens, the molecular weight in equation (19.4) gradually shifts from M_w to the next higher average, M_z. Viscosity versus temperature data are available in the literature.[52,54,55]

Melt viscosity also is a function of pressure,[52] but the change is negligible until one operates in the range of 5000 to 25000 psi, where viscosity increases by factors of 10 to as much as 500 can occur.[53] Because static pressures in this range are encountered with injection molding and some extrusion operations, the effect cannot always be ignored.

Polymer melts and solids can exhibit combinations of two types of behavior when deformed: (1) elastic, where the deformation energy is stored as potential energy and is, ideally, fully and instantly recoverable; and (2) viscous, where the deformation energy is not stored but is dissipated over time as heat. These materials are called viscoelastic, and, in analyzing their behavior, both types of deformation must be taken into account.

The degree to which a viscoelastic material behaves as an elastic solid or a viscous liquid depends largely on the time scale and pattern of the applied stress and the response time required by the system. For example, a ball of Silly Putty dropped to the floor will bounce elastically as if it were made of rubber. Here there is a very brief applied compressive stress as the ball hits the floor. If that same ball is allowed to rest on the floor for a sufficient length of time, it will slowly start to flow and end up as a puddle, like a viscous liquid.

Depending on composition, structure, and temperature, a viscoelastic material has a characteristic relaxation time, λ, defined as the time required for a stress to decay to $1/e$ of its elastic response to stopping any change in strain.[56] The response time can be expressed as:

$$\lambda = \eta/G \quad (19.5)$$

where G is the modulus of the material. Figures 19.10 through 19.13 show generalized curves of modulus versus temperature for polymeric materials. There is a high-modulus glassy state that drops off to a rubbery plateau for uncured or lightly cross-linked polymers. T_g occurs at the inflection point of the dropoff. More heavily cross-linked polymers have a higher plateau region above their T_g's, whereas

there is a further dropoff of G for uncured, linear polymers beyond the rubbery plateau.

The ratio λ/t_e, where t_e is the time scale of the experiment, is called the Deborah number.[57] There is elastic response where the Deborah number is high, as with the bounced Silly Putty (t_e is the very brief ground contact time of the bounce.) Where the Deborah number is low, the behavior is viscous (as with the protracted experiment, with Silly Putty flowing on the floor).

Figure 19.19 shows schematically the pattern of strain response to an applied tensile stress with time t for a viscoelastic material. Following application of the stress at t_1, there is an initial, rapid elastic deformation, producing a strain D. There ensues a nonlinear, strain versus time period, following which the relationship becomes virtually linear. This latter pattern, known as creep, can, in principle, be observed with all solids if the time scale is long enough. At t_2 the stress is released, and the material initially contracts elastically in a manner comparable to the initial strain D. Following this, the strain versus time relationship approaches an asymptotic value of nonrecoverable strain, essentially equal to that developed during the creep period. This effect is commonly observed with ordinary rubber bands, which, after weeks or months of remaining stretched, will no longer fully return to their original length when released. The molecular chains slowly and irreversibly began sliding by one another during the extended stretch period instead of behaving as end-anchored, parallel coiled springs. Light cross-linking (i.e., vulcanization or curing) of the rubber inhibits this chain sliding effect. As cross-linking is increased, however, chain mobility will drop to a point where the desired rubbery behavior is lost.

Fig. 19.19. Idealized viscoelastic material response (Burger's model) to constant, protracted tensile stress starting at t_1, followed by sudden release of the stress at t_2.

Many other viscoelastic effects are encountered in the manufacture, fabrication, and use of plastics. Extrudate swell as a melt emerges from a die is a form of elastic recovery from the extensional stresses applied as the melt passes through the die orifice. This must be taken into account with many extrusion operations. Shrink films are made by biaxial stretching at temperatures slightly above T_g, followed by quenching. The resulting built-in film strain is relieved at an exceedingly slow rate at ordinary temperatures because of the very long mean relaxation time. With sufficient heating, however, the relaxation time will drop by orders of magnitude. The built-in strain then rapidly diminishes, and the film shrinks.

A final example is the Weissenberg effect,[58] where a viscoelastic fluid will flow at right angles to the direction of a steady-state shear stress. Thus, materials such as polymer solutions or flour dough will climb up an immersed rotating agitator shaft.

FABRICATION OF PLASTICS

The variety of means and the relative ease by which plastics can be formed into useful objects are important advantages in comparing them with traditional materials. We will survey here some of the principal fabrication techniques employed today. Although our fundamental understanding of many of these operations has broadened considerably, especially since about 1950, a great deal of art remains. In addition, the field is dynamic,

with important design and control innovations continually being introduced.

Extrusion

Extrusion is one of the basic methods of fabrication, particularly for thermoplastics. It consists of shaping a material by forcing it through a die, originally via a hydraulic ram. By the 1930s screw extruders, adapted from rubber technology, were replacing older designs because they could be run continuously, were adaptable to a much broader range of materials than the older equipment, and were capable of vastly superior performance.

In addition to their material-shaping function, screw extruders can be designed for such varied tasks as mixing and blending, addition or removal of volatiles, expelling insoluble liquids, and carrying out controlled chemical reactions. The extrudate might be in the form of pellets of a desired size and shape, usually intended for a subsequent fabrication step. It also might be formed into pipe, tubing, sheeting, film, and profiles (such as weather-stripping and house siding), or be coated directly over wire, cable, and cord, or onto paper. In addition, extrusion is incorporated into other types of fabrication procedures such as injection molding and blow molding, which will be described later.

The workhorse design for plastics extruders is based on a single screw rotating in a horizontal, cylindrical barrel, usually equipped with a hardened liner. A schematic diagram is shown in Fig. 19.20. The extruder barrel generally is equipped with temperature-controlled sections along its length. Heating can be via electricity or heat transfer fluids, with cooling via forced air, water, or other coolant fluids. The screw often is cored for coolant circulation along a portion of its length. The feed material most often flows via gravity from a funnel-type hopper through a feed throat in the barrel and into the channel of the screw. The screw is driven by a motor through a gear reducer with a thrust bearing positioned to absorb the rearward thrust of the screw. The feed melts as it is conveyed along the screw channel, and at the downstream end of the barrel the melt is forced

Fig. 19.20. Elements of a single-screw extruder. (*Bernhardt, E. C.,* Processing of Thermoplastic Materials, *copyright by the Society of Plastics Engineers, Inc., 1959. Van Nostrand Reinhold, New York.*)

Fig. 19.21. Single-stage extruder screw. (*Courtesy Spirex Corp.*)

through a breaker plate, which frequently supports a screen pack to filter out contaminants. From there the melt exits through a shaping die. Single-screw extruders usually are sized on the basis of the inside barrel diameter, which can vary from as small as 0.5 inch for special laboratory or industrial applications with capacities down to a few pounds per hour, to as large as 24 inches (600 mm) and capacities of over 60,000 lb/hour. Most industrial units are in the 1.5 to 8-inch range.

The heart of the extruder is the screw, and the success of an extrusion operation is largely dependent on its design. Since the early 1950s, a large volume of literature has been published on the theoretical and practical aspects of extruder screw design and operation.[58,64,92–95] A typical single-stage extruder screw is shown in Fig. 19.21. The screw helix is very commonly "square pitch"; that is, it advances one turn per unit length equal to the screw diameter. This is equivalent to a helix angle of 17.7° at the outside of the screw. The principal screw variable with this design is the change in root diameter with length. The screw, as illustrated, usually has three sections: a relatively deep-flighted feed section, followed by a tapered transition section, and, finally, a shallow-flighted metering section. The length and flight depth of each are determined largely by the frictional and melting behavior of the feed, the rheological properties of the melt, the pressure drop of the screen pack and die, and the desired output rate. Melt thermal stability and special mixing or blending needs also play important roles in this screw design. The clearance between the screw flights and barrel must be tight enough for efficient melting and pumping and good heat transfer, but loose enough to avoid binding. The radial clearance commonly employed is about 0.1 percent of the barrel diameter.

With a solid feed to the extruder, such as pellets, powder, beads, flakes, and regrind, alone or in combination, most of the energy for melting comes from friction and viscous dissipation rather than heat transfer through the barrel wall. The melting of feed in the screw channel, a complex process, is illustrated schematically in Fig. 19.22. With these designs, there is generally little heating or melting in the feed zone of the screw (1). However, as the feed enters the transition zone, it starts to be compacted because of the decreasing channel depth and is forced against the barrel, causing a sliding friction that creates a film

Fig. 19.22. Schematic of feed melting in a conventional single extruder screw channel. (*Courtesy Spirex Corp.*)

of molten polymer there (2). This molten material is collected as a melt pool ahead of the advancing flight (3). The molten polymer circulates in a helical path. As the material continues along the transition zone, this melting continues. The melt pool width increases, and the solid bed shrinks (4). At some point the diminished solids bed often will break up, leaving the remaining solid particles dispersed in the melt pool (5 and 6). Because the high shear stresses previously developed against the barrel no longer are possible, the remaining melting is primarily via thermal conduction from the melt, which is far less efficient than previous melting. Any remaining solids generally will be melted in the metering section where shear rates are higher.

Although there are many complexities in this process that are beyond the scope of this chapter, the output characteristics of a properly designed square-pitch screw can be considered to be largely functions of the diameter, length, and flight depth of the metering section. In addition to output, the design of this section is of major importance in setting melt temperature and flow uniformity. The shallow-flighted metering section has been most tractable to theory because geometric factors there are relatively simple, and melt properties can be most readily defined. In its most basic form, the flow from the metering section, Q, is made up of three components: drag flow, Q_d, pressure flow, Q_p, and leakage flow, Q_l.

Drag flow is simply the forward conveying action developed by the relative motion between the screw and the barrel in the absence of any downstream flow restrictions such as a screen pack or die. With shallow channels and a single-flighted screw of narrow flight width:

$$Q_d = (\pi DhS \cos^2 \phi)N/2 \qquad (19.6)$$

where D = screw diameter, h = flight depth, S = axial channel width, N = screw revolutions per unit time, and ϕ = screw helix angle = $\arctan(S/\pi D)$. Equation (19.6) also can be expressed as:

$$Q_d = \alpha N$$

where α is dependent only on screw dimensions.

Pressure flow is caused by pressure in the head of the extruder. If we assume that the screw is stationary, but that there is melt under pressure at the die, the screw channel will act like a long rectangular orifice. The melt, forced by the pressure in the die head, will travel backward down the helical screw channel. In actual operation, pressure flow is a reaction to drag flow, caused by the die restriction and the pressure developed in the die head. It is influenced by screw channel dimensions, barrel diameter, melt viscosity, and back pressure. If viscosity is assumed constant, then:

$$Q_p = (h^3 S \sin \phi \cos \phi)(dp/dl)/12\eta \qquad (19.7)$$

where l = axial channel length, p = melt pressure, and η = melt viscosity. Here:

$$Q_p = (\beta/\eta)dp/dl$$

where β depends only on screw dimensions.

Leakage flow is a backflow in the narrow clearance between the screw flight and the barrel caused by the pressure developed in the die head. It is usually negligible with single-screw extruders except where there has been significant screw wear.

The net volumetric output of a metering section as described above can be expressed as:

$$Q = Q_d - Q_p = \alpha N - (\beta/\eta)dp/dl \qquad (19.8)$$

Figure 19.23 illustrates how output varies with head pressure according to equation

Fig. 19.23. Characteristic screw and die operating curves. (*Courtesy Monsanto.*)

(19.8), with different metering sections operating under conditions of equivalent screw speed and melt viscosity, assuming newtonian behavior and isothermal conditions. Three metering section designs are compared: (1) a short, deep metering section; (2) a long, deep metering section; and (3) a shallow metering section. The output value at zero head pressure is the drag flow, and the figure shows how this is increased by deeper screw flights. Comparison of the slopes of these lines shows that the output of a short, deep-flighted section is more sensitive to head pressure than a longer section of the same flight depth. The least sensitive configuration illustrated is that where the flight depth is shallow. Because it is generally desirable for output to be relatively unaffected by head pressure, designers will try to provide a long, shallow-flighted metering section. The metering section length may be limited by length requirements of the other screw sections and the overall length of the barrel. If the metering flight depth is too shallow, it may not be possible to obtain the necessary output without excessive shear; this could result in mechanical degradation or overheating of the melt. Also shown in Fig. 19.23 are the output versus head pressure characteristics of two dies: one with a large opening and one with a small opening. Both lines pass through the origin. The intersections of the screw and die operating lines represent the operating points of those particular screw–die combinations.

It frequently is necessary to remove volatiles such as air, moisture, or solvents from the melt before it emerges from the die. For this purpose a two-stage screw with a vented barrel, as illustrated in Fig. 19.24, commonly is used as long as the level of volatiles is relatively low. Here there is an abrupt decrease in root diameter following the first metering section just upstream of the barrel vent. This increase in flight depth is intended to result in partial fillage of the melt channel in that zone, thereby preventing any melt discharge from the vent. The vent thus may be open to atmospheric pressure or connected to a vacuum system, depending on the volatiles in question and the degree of removal desired. The vent zone length with its partial fillage provides residence time for release of volatiles. Following the vent zone, the melt passes in turn through a second transition zone and a metering zone before emerging from the die. The second screw stage must be designed for a higher output rate than the first stage to avoid melt flow from the vent, but not high enough to cause surging or other flow instabilities.

The L/D ratio of screws will vary from about 20:1 to 36:1 or greater, depending primarily on melting, metering, and venting requirements.

The single-screw extruder screws described above have relatively limited capability for mixing and blending; so numerous devices have been developed in attempts to improve on them—generally short sections with L/D ratios of about 2:1 to 5:1, and often attached at the discharge end of the screw. Two types are shown in Fig. 19.25. The Dulmage section has shallow, multiple flights with semicircular melt channels usually separated by three or more short, undercut cylindrical sections that interrupt the laminar flow, dividing and recombining the melt many times. It often is used with foam extruders. The Union Carbide (Maddock) section features a series of semicircular grooves along the screw axis, of which

Fig. 19.24. Two-stage screw to be used with a vented barrel. (*Courtesy Spirex Corp.*)

Fig. 19.25. Single-screw mixing/blending sections. (*Courtesy Spirex Corp.*)

alternate grooves are open to upstream entry, while the others are open to downstream discharge. The mixing lands shown are slightly undercut to provide a greater clearance with the barrel, compared with the slightly narrower wiping lands. The polymer is pumped to the inlet groove and is subjected to a brief, high shear as it passes over the undercut mixing land. Then it is pumped out of the discharge groove because it cannot escape over the tight-clearance wiping land. This arrangement screens out unmelted materials and thus can be designed with deeper channels than other designs have, for greater output with equivalent or better mixing. It is used most often with polyolefins, especially low density polyethylene.

Barrier and other duel channel screws have been developed in recent years, primarily to improve melting efficiency, mixing, and output. There are numerous designs,[59,60] generally functioning as a transition section, most of which feature a secondary or barrier flight, which creates two distinct channels, one for the melt and one primarily for the solids. The principle is illustrated schematically in Fig. 19.26. The barrier flight has greater clearance with the barrel than the primary flight. This allows melted plastic to flow over it while blocking the passage of feed solids. Melting in these barrier screw sections proceeds in a manner similar to that described earlier, but the channel for melt gradually enlarges in the downstream direction while the solids channel gradually becomes narrower and/or shallower. This decrease in size of the solids channel prevents breakup of the solids bed, which frequently occurs with conventional screws as shown in Fig. 19.22. Melting therefore can continue by the efficient mechanism of frictional heating against the barrel. The Double Wave screw,[65] shown in

Fig. **19.26.** Schematic of feed melting in the channels of a barrier type screw. (*Courtesy Spirex Corp.*)

Fig. 19.27. Double Wave screw. (*Courtesy Spirex Corp.*)

Fig. 19.27, is a variant generally employed as an extended metering section. It has two channels of equal width separated by an undercut barrier flight. The roots of each channel go up and down in wave fashion, such that the depth of one channel is shallow while the channel across the barrier is deep. The melt is thereby forced back and forth across the barrier, and is subjected to alternately high and low shear. Any solids particles are wedged into each wave crest, where they are briefly exposed to high shear. These features promote melting and mixing, and high output can be achieved at lower melt temperatures with improved pressure stability.

A feature increasingly incorporated into some single-screw extruders is the grooved barrel.[61] With this design there are longitudinal grooves in the barrel extending 3 to 5 diameters from the feed port, with their depth decreasing to zero in linear fashion. These grooves greatly increase the frictional heat generation at the barrel surface, so the solids conveying zone can develop very effective conveying characteristics with resultant high pressures. The extruder output thus becomes controlled by the feed zone rather than by the plasticating and melt conveying zones. The results are reported to be a greater and more stable output and reduced sensitivity of the output to pressure. These advantages are at least partially offset by lower efficiency (i.e., kg/(kWh)), greater wear, and increased design complexity.

Twin-screw extruders are more expensive than single-screw types, but they have been widely employed with hard-to-feed materials, where compounding and devolatilization requirements are more critical, where melt thermal stability is limited, or where chemical reactions also are being carried out. There are several design concepts, depending on whether the screws are corotating or counterrotating, and the degree to which they intermesh. These systems and variants are illustrated in Fig. 19.28.

With counterrotating extruders, the screws at any common point have the same helix angle but are oppositely pitched. Some intermeshing types feature conical rather than parallel, constant-diameter screws. The large diameter of such screws at the feed end promotes heat transfer and allows greater

Fig. 19.28. Twin-screw systems. (*Courtesy Werner & Pfleiderer Corp.*)

shaft separation, which provides the room needed for more rugged gears and thrust bearings. Such screws also can be designed for higher torque, which facilitates running at low speeds, thereby helping to prevent excessive melt temperature. Screw rotation is generally outward at the top and inward at the bottom. Because the linear speeds at the metal nip surfaces are unequal except where the channel depths are the same, there is pronounced calendering of the melt. This, along with the relatively slow screw speed, positive conveying behavior, short residence time, and narrow residence time distribution, favors such applications as the extrusion of unplasticized PVC which is thermally sensitive, as well as complex profiles.

The screws of nonintermeshing counterrotating extruders rotate opposite to the intermeshing type described above, that is, inward at the top and outward at the bottom. This favors feed acceptance and the sequential feeding of separate solid or liquid ingredients desired in some compounding and reactive extrusion operations. Such an extruder is illustrated in Fig. 19.29. Because there is significant clearance at the top and bottom apexes where the cylindrical barrel segments come together, there is leakage flow which increases distributive mixing—a desirable feature for some applications although detracting from output. There is also considerable exchange of material between the screw channels, which favors mixing and blending. These extruders frequently are employed for incorporating glass fibers into melts because of their excellent feed acceptance and capability for gentle mixing, which helps minimize breakage. The machines can operate efficiently at partial fillage over considerable screw lengths. This, along with the inward drawing feature at the top apex, makes the design very effective for devolatilization.

The screw design often will involve short sections with different pumping, mixing, and pressure development characteristics, which can be assembled in modular fashion with appropriate barrel sections to accomplish a variety of functions such as those outlined above. They can occasionally reach L/D ratios as high as 60:1 or more. As a result of the leakage flows mentioned above, these extruders have diminished capability to force the melt discharge through some die restrictions. The needed pressure generation can be achieved by feeding the discharge to a close-clearance melt gear pump or into a single-screw extruder. In the latter case, it is often accomplished by appropriately extending one of the two screws and its barrel beyond the downstream end of the other, as shown in Fig. 19.29.

The screws of intermeshing corotating extruders are generally designed to be self-

Fig. 19.29. Counterrotating, nonintermeshing twin-screw extruder. (*Courtesy Welding Engineers, Inc.*)

wiping, and, along with appropriate barrel sections, they can be assembled to perform many different functions such as feeding, conveying, mixing, kneading, venting, and pressurizing. Individual sections generally are heated and cooled separately. The feed, in any variety of solid or liquid forms, is metered rather than hopper-fed as with most single-screw machines. The corotating feature conveys most feeds very efficiently. Some of these screw configurations are shown in Fig. 19.30. Forwarding and pumping screw sections might have up to three flights, resulting in different cross sections, as shown in Fig. 19.31. Unlike the intermeshing, counterrotating extruder, there is no calendering action between the screws with the corotating design; so they are not pushed apart as they rotate. This leads to reduced wear against the barrel and permits higher screw speeds. These machines also can be effective devolatilizers and reactors because virtually all the material is transferred from one screw to the other with each revolution, resulting in complete renewal of surfaces, and sections are intended to operate at partial fillage.

Computer simulations have been developed that are of considerable assistance in screw design, especially with single-screw

Fig. 19.31. Cross sections on intermeshing, corotating screw elements with one, two, and three flights. (*Courtesy Werner and Pfleiderer Corp.*)

Fig. 19.30. Simple screw elements for a corotating twin-screw extruder. The screw elements are assembled onto keyed shafts. (*Courtesy Werner & Pfleiderer Corp.*)

machines.[62] They go far beyond equations (19.6) through (19.8), discussed earlier. The melting step is generally the most challenging in developing a workable model because the onset of melting and conditions leading to breakup of the solid bed are difficult to formulate reliably. Simulations are of particular value in assessing the effect of a process variable change, but they rarely are relied on entirely in the design of a commercial unit.

The extruder die shapes the melt stream leaving the extruder. Following the extruder is equipment designed to hold or further modify the shape of the extrudate and to cool it until it is hardened. The extruder discharge must be homogeneous, nonsurging, and of uniform temperature, particularly where dimensional tolerances of the final product are tight.

The relationship between flow and pressure through a die is given by:

$$P = K\eta Q \qquad (19.9)$$

where Q = volumetric flow rate, η = melt viscosity, and K = die resistance. Die resistance is a function of the length and cross-sectional dimensions of the die, which can be very complex with some designs. Because most polymer melts are nonnewtonian, the viscosity will depend on shear rate, which, in turn, is a function of local channel dimensions. Although die resistance and shear rate formulas are available for relatively simple geometries,[66] the actual application of equation (19.9) generally is carried out with computer models that proceed incrementally along the flow channel, adjusting the viscosity to the shear rate and temperature in each increment. The aim is to provide the necessary production rate at an acceptable pressure and temperature, and to obtain a finished product of the required shape and dimensions.

Dies come in many designs depending on the product: sheet, cast film, blowfilm, pipe, profiles, pellets, and so on. They often are equipped with means to adjust flow channel dimensions at sensitive points (e.g., along the lip gap in sheet and film dies) in order to equalize flow rates across the die and satisfy critical dimensional specifications. Uniform heating of the die also is of great importance, for similar reasons. In making pellets, the die forms strands like spaghetti, which either are cut by rotating knives at the die face and then cooled, or are drawn through a cooling water bath and then passed to a pelletizer. Dies for coating wire are side-fed with the extruder at an angle of 30° to 90° to the running direction of the wire. The melt coats the wire as it exits a mandrel enveloped by the melt. This is illustrated in Fig. 19.32.

A pipe die provides a carefully sized annular opening for the melt, which then is pulled through a long water trough where it cools. A common sheet die design is shown in Fig. 19.33. The melt from the extruder, which is

Fig. 19.32. Basic design of a wire coating die. (*Reprinted from* Extrusion Dies, *W. Michaeli, by courtesy of Carl Hanswer Verlag, 8 Munchen 80, Vilshafener Str. 10.*)

Fig. 19.33. Slit die for flat film or sheet extrusion. (*Reprinted from* Extrusion Dies, *W. Michaeli, by courtesy of Carl Hanser Verlag, 8 Munchen 80, Vilshafener Str. 10.*)

of circular cross section, is spread out into a rectangular sheet via a "coat hanger" or similar manifold, followed by a land, or dam, that serves as a flow restriction zone to promote uniform spreading of the melt across the manifold.

The contour and the cross-sectional shape of the manifold, along with its downstream flow channel height, and the way the latter two characteristics vary with lateral distance from the die centerline, are intended ideally to produce uniform melt flow along the entire die lip so as to form an extrudate sheet or film of constant thickness. Because this normally cannot be fully achieved in practice, dies are designed to permit changes in flow channel dimensions where needed. Two techniques commonly are employed. A choker bar, as shown in Fig. 19.33, can be locally deformed by a series of bolts along the width of the die, thereby changing the height of a portion of the flow channel. In addition, the die gap is designed to be locally adjustable by means of a series of compression bolts along the upper die lip across the width of the die. This "flex lip" feature is usually the principal means of adjustment. With some designs, the bolts are expandable via adjustable individual electrical heaters. This means of varying bolt compression of the flex lip also can be the basis of a system for automatic control of sheet gage. The lower die lip usually is set at one position, depending on the nominal thickness of the product. The hot sheet exiting such dies usually passes directly to a series of two or more carefully gapped, temperature-controlled, chrome-plated rolls, where it is cooled and the final thickness and degree of surface gloss are imparted, before it is edge-trimmed and cut into lengths or wound into rolls.

A great deal of plastic film, especially polyethylene, is made by a blow-film process, illustrated in Fig. 19.34. Here the melt passes through a circular die with an adjustable, narrow annular opening. As air blowing from an encircling ring cools the hot melt exiting the die, compressed air admitted through the center of the die expands the thin-walled extrudate into an elongated balloon. The cooled exit end of the balloon is closed by a pair of nip rolls. The flattened tube then can be rolled up as shown for subsequent fabrication into bags. Alternatively, it can be edge-trimmed, with the two layers separated and then slit appropriately to make film stock as wide as 40 feet.

As the melt is cooling, it is stretched in two directions: circumferentially with the growing diameter of the balloon, and longitudinally as it is pulled towards the nip rolls. When the temperature range and degree of biaxial stretching are properly set, the polymer molecules will tend to become oriented along the directions of stretch rather than in their otherwise generally random pattern. The result will be films with enhanced strength and toughness.

Fig. 19.34. Formation of sheet by extrusion through a circular die (blown bubble extrusion). Ingredients are resins, stabilizers, pigments, and other additives. (Encyclopedia of Polymer Science and Technology, *Vol. 6, H. G. Mark and W. G. Gaylord, Eds., Interscience Publishers, 1968; copyright © 1968 by John Wiley & Sons, Inc. and used by permission of the copyright owner.)*

Coextrusion is an important technique for producing a well-bonded, multilayer product as sheet, film, tubing, wire coating, and most of the other forms made by conventional extrusion. Such structures can be made to combine the properties of the resin used in each layer (e.g., gas barrier capability, moisture resistance, lower cost, chemical resistance, etc.). Because melt flow downstream of the extruder screw is invariably laminar, if two or more polymer streams can be combined as layers in a suitable die system, it is possible for them to emerge from the die in the same layer relationship. Coextrusions have been made containing as many as nine or more layers.

Each component melt, which can constitute one or more layers in the final product, is fed to the die system from its own extruder. Some die systems employ a feed block where the melt streams are layered in a relatively small cross section before entering the die, which is of basically conventional design. This assembly is the lowest cost but permits little melt flow variability among the individual components, and lacks layer thickness control. With another design, there is no feed block. Instead, the melt streams enter the die separately and are brought together just upstream of the die orifice. Here, a greater degree of flow variability and thickness control is possible. This design is more costly than the first and is difficult to use with more than three layers. In a third technique, used primarily with sheet and film, each melt stream has its separate manifold and exit orifice. The molten layers are combined just downstream of the die exit. With flat film dies this is accomplished with pressure rolls, whereas with blow-film dies the layers are combined by the air pressure within the expanding bubble. This technique reportedly provides accurate control of layer thickness and is easier to start up than the systems described above.[63]

Injection Molding

Injection molding is the most commonly used method for converting thermoplastic materials into final products, and it is an important and growing technique with thermosetting resins as well. Here molten plastic in a cylinder is pushed out under high pressure, by either a piston or an extruder screw acting as a piston, through a nozzle that abuts firmly against an opening into a relatively cool, closed, two-piece mold. The plastic then solidifies to form a part shaped like the reverse image of the mold. Then the mold is separated, the part is removed, and the cycle can be repeated. Injection molding equipment can be automated and is capable of high, economical production rates with relatively low energy

usage. Parts can be made to close tolerances in a broad range of surface textures, with little or no finishing required. In many cases parts can be made in a quantity that would be very difficult or impossible to achieve by other methods. Scrap losses tend to be low because rejects, runners, and gates can be directly reused after regrinding. Combinations of plastics and fillers can be molded, and it also is possible to incorporate metallic and nonmetallic inserts during the molding step. Among its disadvantages are expensive molds and equipment, difficult process control, and inconsistent operation.[67]

Reciprocating screw designs make up most of the injection molding machines used today. The cyclic operation of such a machine is shown schematically in Fig. 19.35. At the beginning of the cycle, shown in Figure 19.35A, hydraulic pressure acting on the clamp ram toggle (1) forces the moving platen containing one part of the mold (2) against its mating portion mounted on the stationary platen (3), thus closing the mold (which is cored for circulation of cooling water.) In Fig. 19.35B the hydraulic cylinder (4) brings the injection assembly (5) forward into sealing contact with the sprue bushing (6) of the now closed mold. In the next step, shown in Fig. 19.35C, the hydraulic cylinder (7) forces the nonrotating screw forward to inject plastic melt into the mold cavity. A valve at the screw tip prevents melt from leaking back into the screw flights during this step. The hydraulic oil velocity and pressure moving the screw are set to fill the mold with the plastic quickly and completely, and to ensure against defects such as sink marks and welds. The force exerted by the clamp ram toggle is kept greater than that exerted by the flowing plastic to ensure that the two mold sections stay firmly together. If they did not, there could be melt leakage from the mold, producing a defect called flash, which usually must be removed and often results in part rejection. In Fig. 19.35D the melt cools in the mold and solidifies to the point where the injection pressure exerted by cylinder (7) can be released. The screw then starts to rotate and charges melt to the front end of its cylinder (9). The screw tip valve mentioned earlier opens when rotation starts, permitting the melt to flow in front of it and force the screw back. In the next step, shown in Fig. 19.35E, the injection assembly moves away from the sprue bushing via the hydraulic cylinder (4). The screw still is charging the front end of the cylinder while moving back in its barrel. While the screw is turning, plastic feed, generally in the form of pellets, drops from the feed hopper (10) into the screw channel. Melting of the feed is accomplished in much the same way as in conventional extrusion: via the mechanical energy expended in turning the screw and, usually to a lesser extent, by contact with the electrically heated barrel. Screw rotation stops when the rearward motion of the injection assembly touches an adjustable limit switch. Its location determines the amount of material remaining in front of the screw, which is the size of the shot. Figure 19.35F shows the final step of the cycle, where the clamp ram withdraws the movable platen which opens the mold and allows the part (11) with its sprue to be ejected.

Because of the high injection pressures required to ensure that viscous melt properly fills the mold in the requisite time, and the resultant high clamping force needed to keep the mold closed during the injection step, the entire assembly must be ruggedly built, yet capable of precise, repeatable operation. A good approximation for design purposes is 2.5 tons of clamping force for every square inch of projected mold area.[68] Clamping force, in tons, is one of the main specifications for injection molding machines, and it varies from 5.5 tons for machines with 0.37 cu. in. injection capacity to 6050-ton machines with 6000-cu. in. injection capacity.

The diameter of the screw is largely determined by torque limitations. Peripheral screw speeds are usually limited to about 110 ft/min or less to avoid shear degradation of the polymer. The higher power required for higher output will therefore call for larger diameter screws to prevent breakage. (The torque rating of a screw is proportional to the cube of its diameter.) With a 2.5-inch diameter screw, for example, the maximum permissible

A
B
C
D
E
F

speed is 200 rpm. With most screw designs the maximum drive input will be about 40 hp.

Reaction Injection Molding (RIM)

With the RIM process, first commercialized in the late 1960s, generally two individually stable streams of low viscosity components of a polymerizing system are injected into a small chamber where they are rapidly mixed together by impingement and then flow from there into a mold where polymerization takes place. When the reaction has reached a satisfactory level of completion the mold is opened and the plastic article removed. The cycle can then be repeated.

The processing requirements for RIM are very stringent: (a) the reactive streams must be effectively mixed by impingement; (b) the system must not produce molding defects due to turbulence or advance polymerization; (c) completion or near-completion of the polymerization must be achieved in an economically acceptable time and without excessive exotherms; and (d) the reaction must be free of undesirable by-products. To date, very few polymerization systems have been used commercially with RIM. Polyurethanes account for over 90% of RIM materials produced.[69] Other RIM systems include the anionic polymerization of caprolactam to produce nylon-6,[70] and the ring opening polymerization of dicyclopentadiene.[71]. Fillers are often incorporated into RIM formulations to lower raw material costs and to improve certain properties such as stiffness.

RIM has certain characteristics in common with injection molding. However, there is a basic distinction in that with RIM a low-viscosity, reactive mixture is injected into the mold rather than a viscous polymer melt. Injection pressures with RIM are therefore quite low, and require only moderate mold clamping pressures. Heavy duty equipment and molds are therefore not necessary, and this lowers capital and tooling costs. Since these savings grow with increasing machine size, RIM has found particular application with large parts with high surface/volume ratios. Another economic advantage of RIM is its very low energy requirements. Unlike most other plastics fabrication methods, the components need little or no heating, and mold temperatures are rarely above 80°C. This is especially true with polyurethanes, but less so with nylon-6. The polymerization exotherms provide most of the heat needed to achieve the needed molding temperatures.

Figure 19.36 is a schematic drawing of a RIM machine. It shows two jacketed, air- or nitrogen-blanketed hold tanks for the reactants. Each tank has an associated positive displacement metering cylinder that feeds to an impingement mixer assembly connected to the mold. The tanks have agitators to maintain uniformity of charge temperature and composition. The metering cylinder pistons are raised via the hydraulic cylinders shown, which draw in reactants from their respective storage tanks. When the metering cylinder pistons are forced down, a series of valves in the lines direct their flows into the mixing head through small orifices in its side wall. This mixing head can operate in two modes, depending on the position of a third hydraulically operated piston. In the recirculation mode, the streams are returned to their respective storage tanks without any contact. In the mixing mode, the two streams

◀

Fig. 19.35. Schematic of injection molding operation. A. When the mold starts to close, the screw has just finished charging the front end of the cylinder. B. With the mold closed, the heating cylinder then moves toward the sprue bushing and is ready to commence the injection cycle. C. The hydraulic cylinder then forces the screw forward under the injection head and fills the cavity. D. Cooling and charging cycle: the mold is filled and, after the holding time, the screw starts to rotate and charge the front end of the cylinder. E. The screw is still charging the front end of the cylinder, and the cylinder starts to move away from the sprue bushing. F. The finished cycle: the mold opens and the molded part is ejected with the sprue bushing. (*Du Bois, J. H., and John, F. W.,* Plastics, *Reinhold, 1967; by permission of Van Nostrand Reinhold Co.*)

Fig. 19.36. Reaction injection molding (RIM) process flow sheet. (*Courtesy Monsanto.*)

impinge at high speed in the mixing chamber at a 180° angle to each other. The turbulence created by this impingement is great enough to avoid the need for mechanical agitators in the mixing chamber. In addition, at the end of the mixing mode, the mixing head piston moves forward and pushes any remaining reactive mixture into the mold. Both of these features eliminate the need for solvent cleaning and purging of the mixing chamber.

Mold fill times usually range from 1 to 10 seconds, except for the larger molds, which take longer. Cycle times with polyurethanes are usually in the range of 30 to 60 seconds. Parts sometimes are molded as a structural foam with a specific gravity of 0.7 to 0.9 to help ensure complete filling of the mold cavity. The exotherm of the polymerization reaction provides the temperature needed to activate the blowing agent usually present in the RIM formulation. The temperature gradient developed by the polymerizing mixture in the mold usually results in a desirable skin-core foam structure.

The principal market for RIM materials has been the automotive industry, for bumpers, fenders, hoods, door panels, and other large parts. Applications have also been found in business machine housings, construction, furniture, and structural parts.

Compression and Transfer Molding

These fabrication techniques are used primarily with thermosetting resins. With both, the molding compound is heated to soften it enough to flow under pressure into a mold cavity. Temperatures around 150°C are typical for the most commonly used materials. The compound then is held in the mold at pressures around 2000 psi or more for a sufficient time to promote cross-linking and polymerization. This results in a hard, rigid product. In many cases the mold charge is preheated to shorten the cure time.

In compression molding the molding powder is placed in the bottom half of an open heated mold mounted on the lower platen of a hydraulically operated press. The top half of the heated mold, mounted on the upper press platen, is brought into contact with its mate

as the press is closed. The molding powder softens and flows under the pressure exerted by the two mold halves, assuming the shape of the mold cavity. When the plastic material has had enough time to set, the press is opened, and the molded part is removed.

Transfer or plunger molding differs from compression molding in that the molding powder is plasticized before it reaches the mold. This process was developed for the molding of intricate parts with small, deep holes or numerous metal inserts. Because it has been plasticized, the material is able to flow around these inserts without forcing them out of position. The molding compound charge, usually preheated by infrared or dielectric systems, is delivered to an open pot on top of the mold. A plunger, usually driven hydraulically, forces the softened compound through a tapered sprue, followed by runners and gates, into the mold cavities. The pressure capability of these machines usually is on the order of 20,000 psi. Following cure, the plunger is withdrawn, the mold is opened, and the parts are ejected. Fabricators often remove the part from the mold before curing is complete in order to shorten the molding cycle. The final cross-linking then is carried out via postcuring, where the parts typically are held for 2 hours at 135°C, followed by an equal time at 65°C.[72] Commercial machines can be highly automated in terms of cycle repetition, closed loop control, degating, and the separation of parts and runners.

Pultrusion

As its coined name implies, this process employs pulling and extrusion. It is used to produce continuous lengths of fiber-reinforced profiles. Most pultrusion products are based on thermosetting polyesters, although epoxy resins, polyurethanes, and vinyl esters also are being used.[73] Rod- and bar-type profiles generally are reinforced with glass fiber roving. Such structures are highly anisotropic, with the longitudinal strength often more than five times greater than the strength in the transverse direction. Larger profiles, however, also utilize glass matting and fabrics to take advantage of the multidirectional fibers and thereby achieve greater strength in the transverse direction.[74] A mat-roving structure is illustrated in Fig. 19.37. Besides glass, reinforcing fibers have also been made from carbon, aramids, polyacrylonitrile, nylon, and polyester.

Figure 19.38 shows a typical pultrusion line. Roving and mats are drawn through a resin impregnation tank that is designed to keep the roving in good horizontal alignment and to remove excess resin from the reinforcement exiting the tank. The preforming fixture, made of fluorocarbon polymer or ultra-high-molecular-weight polyethylene for easy cleaning, or chrome-plated steel for longer wear, consolidates and shapes the impregnated fibers for entry into the heated, chrome-plated mirror-finish die. The finished profile typically

Fig. 19.37. Mat roving structure for pultrusion. (*Reprinted from the* Pultrusion Technology Handbook *with permission of the publisher, Routledge, Chapman and Hall.*)

Fig. 19.38. Pultrusion line. (*Reprinted from the* Pultrusion Technology Handbook *with permission of the publisher, Routledge, Chapman and Hall.*)

is pulled from the die by a pair of continuous caterpillar belts, and then is cut to lengths with a conventional cutoff saw. For hollow shapes a mandrel is installed upstream of the resin tank and extends through the forming die.[73]

Pultruded profiles have excellent strength/weight ratios, corrosion resistance, and good electrical resistivity, and are easy to fabricate via drilling, centerless grinding, and so on. Therefore, they have broad application in products such as fishing rods, industrial ladders, electrical switch actuators, cable trays, fume scrubber structural members, farm animal pen slats and electrical tower insulated guy wires.[73]

Blow Molding

Although the principal application of blow molding has been for bottle fabrication, it is being used increasingly for other parts such as automotive fuel tanks, seat backs, flexible bellows, and toys. High density polyethylene, polypropylene, poly(vinyl chloride) (PVC), and poly(ethylene terephthalate) (PET) are the principal resins molded in this way. For more special applications, however, other thermoplastics such as nylon, polyacetal, ABS, polycarbonate and ultra-high-molecular-weight polyethylene also are used.[75]

There are three basic blow molding techniques used today. With each of them, a heated preform or parison is inflated by air pressure of 25 to 150 psig, against the two generally concave halves of a closed mold.[76] The part is allowed to cool in the mold; and then the mold is opened and the part ejected.

The most commonly used of these techniques is extrusion blow molding, where the parison is a hollow, thick-walled tube of melt that is extruded vertically from a crosshead die. (A schematic diagram is shown in Fig. 19.39.) The parison then is surrounded by a split, cooled mold of the desired shape. One end of the parison is clamped around a blow pin, which incorporates an inlet for compressed air that expands the parison to make it contact the mold wall. The molded object is ejected when solidified, and excess material is trimmed from the ends and recycled. Containers in the range of 0.25 to 5000 liters are made this way, and integral handles as well as special shapes can easily be designed in.

With injection blow molding, the parison is test-tube-shaped and formed by injection molding. The process, shown schematically in Fig. 19.40, is used for small bottles and parts (rarely more than 1 liter in volume), and is virtually scrap-free. However, the benefits of near scrap-free operation with this procedure must be weighed against its higher tooling costs compared to extrusion blow molding. The method provides very accurate control of part weight, but will not make bottles with handles.

The third process, stretch blow molding, utilizes biaxial orientation, which is possible with certain resins such as PET, PVC, polypropylene, and acrylonitrile–styrene copolymers. Because of the improved strength and stiffness, weight reduction is usually

Fig. 19.39. Basic blow molding process: (1) extrusion of parison; (2) compressed air inflates parison; (3) ejection of blow molded container. (*Reprinted from* Blow Molding Handbook, *Rosato and Rosato, Eds., by courtesy of Carl Hanser Verlag.*)

Fig. 19.40. Basic injection blow molding process: (1) injection molding the preform; (2) blow molding and part ejection. (*Reprinted from* Blow Molding Handbook, *Rosato and Rosato, Eds., by courtesy of Carl Hanser Verlag.*)

Fig. 19.41. Basic stretch blow molding process: (1) injection molding the preform; (2) reheating the preform; (3) stretch blow molding and part ejection. (*Reprinted from* Blow Molding Handbook, *Rosato and Rosato, Eds., by courtesy of Carl Hanswer Verlag.*)

possible. Stretch blow molding is limited to simple bottle shapes, generally in the size range between 0.5 and 3 liters. It is illustrated in Fig. 19.41. Here the parison, which can be made either by extrusion or injection molding, is carefully temperature-conditioned to control the degree of crystallinity. The actual blow molding step generally is carried out just above the glass transition temperature, followed by rapid cooling and part separation from the mold. There also will be trim scrap with this process if the parison is made via extrusion.

The parison may be coinjected or coextruded to take advantage of special properties of the various plastic layers, each of which could provide a special property such as gas or solvent barrier, resistance to aggressive media, printability, surface gloss, toughness, and so on. There usually are greater problems with recycling if there is trim scrap from such operations because the layers may not be easily separable and mutually compatible.

To achieve high blow molding production rates, the parisons are produced continuously or intermittently, and from single or multiple die heads operated in conjunction with multiple molds. There are many machine designs for carrying out the subsequent molding steps.[76] For example, the molds may

be mounted along the rim of a step-sequenced vertical wheel or a circular horizontal table. Each hot parison is, in turn, captured by the closing halves of a mold, then inflated with air against the cooled mold walls, and finally ejected from the opened mold when sufficiently solidified. Then each mold is ready in step sequence to repeat the process.

Thermoforming

In thermoforming, a flat thermoplastic sheet is heated to its softening point, then pressed by low pressure against a mold and allowed to cool. The molded article is die-cut from the sheet. The technique is adaptable to virtually any material with enough hot strength to support itself during the softening step, and has some distinct advantages over other fabrication methods. Parts of large surface area and very thin walls can be formed with relatively low capital outlay and tooling cost, primarily because only low pressures are required. Production parts can be run on relatively inexpensive epoxy and aluminum molds, and wood molds can be used for prototype parts. Coextruded or laminated sheet stock can be run to obtain a wide variety of finished product properties. Well-molded parts do not have excessive molded-in stress. With pressure forming techniques, it is possible to achieve detail quality very close to that from injection molding but at much faster cycle times. It is relatively simple to mold large-area parts. In addition to the high production rates, parts often can be molded, filled, decorated, capped, or sealed in one operation. There are some disadvantages as well: Details can be molded on only one side of the part without special matched tooling. Labor costs can be high in some instances because trimming and detailing may be necessary. There is an additional cost for materials because the feed must be in sheet form rather than cheaper pellets. Trim scrap is commonly 25 to 30 percent and can go as high as 80 percent, so reclaiming is crucial.[77]

Numerous thermoforming techniques and associated equipment are available, with the choice depending on part design, material selection, production volume required, quality, and cost.[78,79] A few of the more common, basic types are described below.

Fig. 19.42. Basic thermoforming. Plastic sheet is clamped and heated. Vacuum beneath the sheet draws it down against the mold. The formed part cools against the mold and then is removed. (*Courtesy* Modern Plastics Encyclopedia, *1989.*)

Straight vacuum forming is illustrated in Fig. 19.42. The plastic sheet is clamped to the mold and heated. The vacuum applied within the mold cavity draws the softened sheet against the mold, where it cools. The sheet areas reaching the mold last are thinnest, and this can be an important limitation with deep-drawn parts. With drape molding, shown in Fig. 19.43, the clamped, heated sheet is forced against the mold, creating a seal. Vacuum applied beneath the mold forces the sheet over the male mold. By draping the sheet over the mold, the sheet area contacting the mold remains close to its original thickness.

Fig. 19.43. Thermoforming with draping. The heated sheet is drawn over the mold, creating a seal. Vacuum then is applied, forcing the sheet against the mold. (*Courtesy* Modern Plastics Encyclopedia, *1989.*)

Fig. 19.44. Thermoforming with matched molds. Heated sheet may be clamped over the female mold or draped over the male mold. The mold is closed, forming the sheet between the male and female parts. Mold vents allow trapped air to escape. (*Courtesy* Modern Plastics Encyclopedia, *1989.*)

Matched molds of wood, epoxy, aluminum, and so on, can be used to press the sheet to shape, as illustrated in Fig. 19.44. The clamped, heated sheet is positioned over the female mold. The mold is clamped together, and trapped air can escape through small vents. Fine mold detail reproduction and dimensional accuracy are possible. Figure 19.45 illustrates pressure bubble–plug assist. Here the plastic sheet is heated and sealed across the female mold cavity. Air, introduced into the mold cavity, blows the sheet upward into an evenly stretched bubble, the height of which is controlled via a photocell. A male plug plunges into the sheet, and vacuum is drawn on the cavity when the plug has reached the lowest point of its downward stroke. This completes the sheet-forming step. Sometimes pressure forming air also is used in this process. Because the sheet is prestretched by this method before contacting the mold, part thickness is relatively uniform.

Commercial thermoforming machines are fed either with precut sheet or from rolls, the choice depending largely on the starting sheet thickness. These machines also have means for indexing or positioning the sheet in the heating, forming, and trimming areas. Heating of the sheet most often is done with radiant elements carefully arrayed to achieve temperature uniformity. Ceramic heaters generally are preferred.

With single-station semiautomatic machines, the clamped sheet is passed in and out of the heating section. With fully automatic operation, all machine functions are performed via a sequential step programmer once the cycle has been initiated. In rotary machines the clamped sheet is rotated from station to station as a means of increasing production and providing for other functions such as trimming. The cycle of these machines is controlled by the longest element, usually heating time, which primarily depends on sheet type and thickness. Some rotary machines may have four stations, of which two are for heating, and which result in a shorter cycle time and better heating control.

High production units generally are fed from roll stock, and often coupled directly to

Fig. 19.45. Thermoforming with pressure bubble snapback. The clamped, heated sheet is sealed across a pressure box. Air pressure applied beneath the sheet forms a large bubble. When the bubble reaches the proper height, a plug is forced into it under constant air pressure. When the male plug closes on the pressure box, higher pressure beneath the sheet and vacuum behind the plug create a uniform draw around the plug form. (*Courtesy* Modern Plastics Encyclopedia, *1989*.)

the output of an extrusion line. This can include coextruded sheet. The heat remaining in the sheet from the extrusion step often is utilized to cut down the heating requirement in the thermoformer unit. A chain with sharpened pins often is used to pierce the sheet near the edges to carry it through the forming cycle. The index length or pin chain step movement is governed by hydraulic stepping motors, servodrives, or an air cylinder operating a rack-and-pinion clutch. Trimming in high-volume applications often is carried out separately in a horizontal press, with capacities up to 4000 articles per minute.

Rotational Molding

This technique, also referred to as rotomolding or rotational casting, is used for the fabrication of a wide variety of products including tanks (up to 20,000 gallons or more), portable toilets, battery cases, scrubber housings and garbage containers. Equipment and mold costs are low, and there is considerable design freedom.[80]

In this process the molds or cavities are charged with a predetermined quantity of resin powder (generally around 500 μm particle size) or liquid. The mold halves then are clamped together and rotated biaxially in an oven where the resin melts and densifies into the shape of the mold. Then the molds are slowly cooled by air and water in a cooling chamber, following which they are opened and the finished articles removed.

The most widely used resins for rotomolding are the polyethylenes, particularly the cross-linkable (XLPE) and the linear low density (LLDPE) types. Low cost and ease of fabrication are among their advantages. Other resins commonly used include polycarbonate for clarity and toughness, ethylene vinyl acetate copolymer (EVA) for excellent low temperature impact strength, nylon for toughness and heat resistance, polyvinyl chloride (PVC) for flexibility, paintability, and applicability as a liquid plastisol, and polyesters for thin wall high impact strength. The resins usually are ground to achieve the desired particle size although some, such as nylon,

can be supplied as pellets of sufficiently small size.

The commonest equipment for rotomolding is made in the form of a carousel with stations for an oven, cooling, loading, and unloading. The unit comes with three or four spindles or arms where one or more molds are mounted. The spindles are mounted on a central hub and are driven by variable-speed motors. The oven generally is gas-fired and equipped with blowers to ensure even distribution of heat. The cooling station, which may not be enclosed, operates with forced air and water, frequently in the form of a spray mist.

For large tanks a shuttle system is used, featuring a frame for holding the mold mounted on a movable bed. Motors for turning the mold biaxially are incorporated in the bed, which is on a track that allows the mold and bed to move into and out of an oven. When the heating cycle is completed, the unit is moved on the track to a cooling station while another bed and mold is moved into the oven from the opposite end.

Foamed Plastics

Foamed plastics are used extensively, not only in terms of tonnage, but also with respect to the resin types that are foamed and the different foam structures that are manufactured.

Most plastic foams are formed by dispersing a gas (often referred to as the blowing agent or pneumatogen) in a molten resin. Such additive blowing agents are of two types: physical and chemical. Physical blowing agents are gases or volatile liquids dissolved or dispersed in the melt, usually under pressure. Such materials include CO_2, hydrocarbons such as the butanes and pentanes, chlorofluorocarbons (CFCs), hydrochlorofluorocarbons (HCFCs), fluorocarbons, and nitrogen. When the pressure is reduced or the temperature increased with these mixtures, existing gas bubbles, or else vapor bubbles evolved from the dissolved blowing agent, will expand to form a foam structure. Chemical blowing agents are compounds that usually are dry-blended with the resin before extrusion or molding. They decompose at fabrication temperatures to form a gas, such as CO_2 or nitrogen, which develops the foam once the pressure is released. There are a great many chemical blowing agents available, with the selection for a given application based on price, decomposition temperature range and rate, and compatibility with the resin. In certain cases the blowing agent is produced by reactions of the polymerization components. An important example is that of some urethane foams, discussed above in the section on thermosets.

If bubbles do not exist initially in a foamable system, as, for example, when the foaming agent is dissolved in the polymer, the bubbles must be nucleated. Control of this step sets the number of growing bubbles and is of great importance in achieving the desired bubble or cell size in the final foam. Nucleation is affected by temperature, system pressure, humidity, the pressure exerted by the blowing agent, and its rate of gas generation. Additives such as surfactants and nucleating agents also are of great importance. The latter may be finely divided solids such as colloidal silica or talc that provide a suitable surface on which bubbles might form. Other agents, such as sodium bicarbonate–citric acid mixtures, react at fabrication temperatures and release small gas bubbles, which, in turn, nucleate bubbles of the principal blowing agent. Contaminants in the polymer also may serve as nucleation sites.

Spherical bubble growth is described by the modified Rayleigh equation:

$$\Delta P = 2\gamma/r + (4\eta/r)dr/dt \qquad (19.10)$$

where ΔP = differential pressure between the inside and outside of the bubble, γ = liquid surface tension, r = bubble radius, η = liquid viscosity, and t = time. At equilibrium, where there is no growth, equation (19.10) reduces to

$$\Delta P = 2\gamma/r \qquad (19.11)$$

Bubbles grow by diffusion of blowing agent vapor out of solution from the surrounding polymer, coalescence of smaller bubbles, and expansion due to thermal and pressure effects

as given by the gas laws. In the development of a foam, the gas first forms separate spheroidal bubbles in the resin. However, as the volume fraction of bubbles in the liquid approaches about 0.70, the bubbles gradually become so crowded together that they start to impinge upon one another and assume polyhedral forms, most often with 12 to 14 faces, each face usually having 4 to 6 sides. The polyhedral struts are formed at the intersection of three faces. Where the gas phase is in the form of separate droplets, the term "high-density," or "structural," foam usually is applied. On the other hand, where the gas forms polyhedral cells, the structure generally is termed a "low-density" foam. If the base resin has a density comparable to water (62.4 lb/ft^3 or 1 g/cm^3), low-density foams from such a resin have a density range in practice of about 0.7 to 15 lb/ft^3 (0.011–0.25 g/cm^3). Structural foam densities would be in the range of 19 to 60 lb/ft^3 (0.3–0.95 g/cm^3).

High-density foams with special electrical and mechanical properties can be produced by dispersing rigid, hollow spheres or microballoons in a liquid polymer, forming a mortar-like material that is shaped by casting and cured. These are called syntactic foams. Epoxy and urea-formaldehyde resins and polyesters often are used as the matrix; and the hollow spheres may be of phenolic resin, glass, or silica, ranging in diameter from 30 to 120 μm.[81] Foams also can be produced by leaching a dispersed soluble phase from a matrix. Cellulose sponges, for example, are made by leaching sodium sulfate crystals dispersed in viscose.

Some plastic foams are made by frothing, where a gas is mechanically dispersed in a latex, followed by coagulation, curing, and drying. Foam rubbers often are made in this way.

The dispersed gas in high-density foams usually acts as an inert filler, serving especially to cut costs. Physical properties of such foams such as modulus, tensile, and compressive strength and Poisson's ratio are largely functions of the void fraction, along with bubble orientation effects.[84,85] Small quantities of dispersed gas sometimes are incorporated in the injection molding of thick-walled parts to avoid sink marks. Structural foams with an integral skin and cellular core can be made by forcing a melt containing a gaseous or chemical blowing agent into a mold at pressures of 100 to 600 psi. The product has a characteristic surface swirl caused by cell collapse there. If an expandable mold is used under high pressure, foaming begins as the mold cavity expands, and the resulting surfaces are very smooth because the skin has formed before expansion takes place.[82] Structural foams also may be produced via extrusion with a die that provides a hollow extrudate, forming a solid skin while the remainder expands toward the center of the profile.[86]

With low-density foams, the relationships of physical properties to structure are very complex because of the wide variation of cell structures possible. Polyurethane foams are probably the most versatile of all. They are capable of being made rigid or flexible, and with open or closed cells. Foam rigidity is controlled by the degree of cross-linking, generally based on the type and concentration of trifunctional reactants in the formulation.

Foams can be formed and molded into their final shape, or produced continuously as slab or bun stock that is trimmed and cut into blocks for particular applications. Besides the diisocyanate, polyols, catalysts, and blowing agents, foam formulations contain surfactants to stabilize the foam and promote cell size uniformity, and selective additives such as flame retardants, fillers, chemical stabilizers, and so on.

The ingredients are metered to a mixing chamber and exit through a nozzle. Cell size and structure are controlled by reactant mixer speed and nozzle length, catalyst type and concentration, and surfactants, as well as by finely divided minerals that act as fillers and nucleating agents.[87,88] Tiny nucleation bubbles rapidly form and are enlarged as the blowing agents diffuse into them. The foaming mixture expands and becomes more viscous as polymerization takes place in the liquid phase. The expansion usually is complete within 1 to 3 minutes after mixing, although further polymerization or gelling continues

which strengthens cell structure. With flexible foams, fluid drains from the cell walls during the expansion step, thinning them to the point of rupture. The cell struts, however, have become firm enough to prevent foam collapse. This produces an open-cell foam. In the forming of rigid foams, the cell walls generally do not burst, and the result is a closed-cell structure. Within 2 minutes after full foam expansion, the gelling step usually is complete. Optimum properties develop as the foam further cures over a period that may last from hours to a few days.[83]

Open-cell foams find particular use in acoustic insulation, cushioning, filtration, and packaging. Closed-cell rigid foams are used for buoyancy, thermal insulation, packaging, and construction.

Polystyrene low-density foams are used in large quantity.[89] These foams, which tend to be closed-cell, are made almost entirely with dissolved blowing agents, generally butanes and pentanes or CFCs. Since about 1989, however, the latter have been phased out for environmental reasons,[104] and they have been replaced with HCFCs. (This subject is discussed in further detail in the next section.) Blends containing CO_2 are becoming increasingly important for the same reasons. The cell size is controlled largely by nucleating agents.

Polystyrene foams are made principally by two processes: extrusion and expandable beads (EPS).[84] The extrusion processes utilizes a single-screw, two-stage machine. Polystyrene pellets dry-blended with nucleating agent are fed to the extruder hopper. The blowing agent under pressure is metered through the barrel at a point coinciding with the start of the second stage of the screw where melt pressure will be low. In the second stage, the blowing agent is mixed thoroughly with the melt, and the temperature is reduced by cooling the barrel to increase viscosity and prevent subsequent foam collapse. In some installations a second extruder in tandem is used for this melt-cooling step. The cooled melt then exits the extruder through a die, generally a circular type. The die channel is designed so that the point where melt pressure has dropped enough for bubble nucleation to take place is very close to the die orifice. This prevents bubble coalescence in the die and promotes uniform cell size. The foam, in the form of a thin-walled circular cylinder, expands rapidly on exiting the die and is drawn over a cylindrical sizing mandrel, following which it is slit lengthwise, opened out flat, and rolled up. The sheet may be postexpanded by passing through an oven, a hot water bath, or a steam chest. Sheet can be made in thicknesses of from 0.016 to 0.375 inch and with densities varying from 2 to 10 lb/ft^3 or more. Such foam sheet is thermoformed into packaging for many applications, and utilized, often in laminates, for thermal and acoustical insulation, graphic arts products, and construction purposes. Polystyrene foam can be extruded with a horizontal slit die fitted with an attached frame that postforms the still molten foam into board up to 2 inches or more thick.[91]

The expanded bead process starts with polystyrene beads made by the suspension process.[90] These are steeped in a pneumatogen, usually pentane, to a concentration of 5 to 7 percent. The bead diameter varies from about 0.01 to 0.08 inch, depending upon the application. The beads first are pre-expanded by heating to 95 to 115°C, usually by steam, but also with hot air, hot water, or infrared methods. The pre-expanded beads then are conveyed to a bin for conditioning, where they are allowed to cool, and the pressure within the closed cells in each bead can equilibrate by permeation of atmospheric oxygen and nitrogen. After conditioning, the beads are fed to a mold where they are heated, generally by steam, and where they expand farther and sinter. Then the mold is cooled, and the finished article is removed. Foam densities generally are in the range of 0.8 to 3 lb/ft^3. This technique has been adapted to make foam board via a continuous process. Foam articles from expanded beads are made by using modified blow molding and injection molding machines. The products are used for thermal insulation, flotation, packaging, and other large-volume applications.

Foams from high-density and low-density polyethylene can be produced by extrusion, usually with CFC blowing agents. They are closed-cell, tough, flexible, and resilient. High-

density foams usually are made from high-density polyethylene with a chemical blowing agent.

Stable low-density polyethylene foams can also be produced by cross-linking. Compounded with a cross-linking agent such as dicumyl peroxide and a chemical blowing agent, the resin is extruded at 150°C into sheet. The sheet then passes through an oven at 200°C, where cross-linking and foaming expansion take place. Cross-linking also can be achieved by X-rays or electron beam radiation. This approach is used primarily with thin foams.

PLASTICS AND ENVIRONMENTAL ISSUES

By recent estimates, plastics waste constitutes about 7 percent by weight in landfills (vs. 37% or more for paper products), and about 16 to 18 percent by volume.[96,100,103] Unfortunately, it also has become a highly visible portion of the litter along roads, on beaches, in parks, and in other sensitive areas frequented by the public. Because plastics products generally degrade very slowly, plastics waste has become a growing source of public concern in recent years. Legislation already has been proposed or passed in the United States on the federal, state, and local levels restricting or banning the use of some plastics in applications where there is a perceived problem.[97] This trend is sure to continue.

The technical community has not lagged in developing responses to these challenges. The principal routes being followed are degradation, incineration, and recycling. Another important approach, which involves both consumers and manufacturers, is source reduction. The activity in this area is focused largely on weaning consumers from the throwaway habits that have developed over recent decades, particularly with packaging and other single-use items that contribute heavily to the plastics waste stream. These efforts are evolving so rapidly that it is difficult to predict how each one will impact the problem over the longer term.

Degradation processes seek to reduce waste plastic to environmentally innocuous materials.[98] One route, biodegradation, accomplishes this through the action of microorganisms that break down the polymer molecules themselves, or else attack a susceptible filler material such as starch. In the latter case, the base resin generally is not profoundly changed, but its macrostructure is broken down so that it becomes unrecognizable. Some systems also employ autooxidant additives that form peroxides when contacted by metal salts in soil or water. These peroxides initiate the breakdown of the polymer chain, reducing its molecular weight so that the polymer can be metabolized by microorganisms, yielding final products that are predominantly carbon dioxide and water. There also was much activity during the 1980s aimed at utilizing microorganisms to synthesize biodegradable polymers for plastics applications. Among the better known of these materials is polyhydroxybutyrate-valerate (PHBV), formed by the action of the soil bacterium *Alcaligenes eutrophus* on sugars obtained from the fermentation of wheat or cornstarch. The plastic is reported to have physical properties similar to those of polypropylene.

Biodegradable plastics are most effectively treated in composting systems where aerobic processes predominate. In general, disposal via modern, sanitary landfills is not very effective with these materials because the limited availability there of oxygen and moisture retards most biodegradation processes.

Photodegradable plastics incorporate photosensitive components to start the degradation process. Ultraviolet radiation in sunlight initiates free radicals in susceptible polymers, leading to chain scission reactions. Because the polymer fragments thus produced usually are brittle and lack physical strength, the plastic eventually disintegrates. There are several ways to accomplish this. For example, polymers with ketone carbonyl systems are easily photodegraded; so ketone-containing monomers are copolymerized with ethylene, styrene, or other monomers, thereby incorporating ketone carbonyls into the polymer chain. When exposed to sunlight, UV radiation is absorbed, causing chain cleavage via

Norrish-type reactions:[97]

Proprietary additives have been developed as photodegradation sensitizers. They include ferric compounds, benzophenones, and mixtures of benzophenones and organometallic complexes. One system, based on iron and copper additives, becomes an active light sensitizer after an initial period of antioxidant and light stabilizing behavior.

With some systems, relatively brief light exposure is enough to start the breakdown, which can continue in darkness. These methods must be carefully tailored to the application so that the normal service life of the plastic article remains unaffected. There is also concern that such materials, if introduced into a recycled commingled plastics feed stream, could have deleterious effects on the stability of the resulting product during normal use.

At present there is a growing consensus that the concept of degradable plastics has been oversold as a solution to the waste disposal problem, primarily because a large portion of degradable plastics will end up in landfills where breakdown tends to be very slow.[96,97,99] The most promising applications of degradable plastics probably are for more limited problems, such as litter, where sunlight, air, moisture, and microorganisms are generally available to accelerate polymer breakdown.

High-temperature incineration of waste plastics has at least two potential advantages. First, increasingly scarce landfill space is conserved because the volume reduction from feed to ash with a well-operated incinerator is in the 90 percent range or higher. Next, the high fuel value of most plastic products can be reclaimed this way for generation of steam, electricity, or both.[100]

Incinerators have drawbacks, however. They can emit corrosives such as HCl and traces of highly toxic dioxins and furans if chlorine-containing materials such as PVC or bleached paper are in the feed stream. These emissions could pose hazards, especially to people living close to the incinerator. Compounds of toxic heavy metals such as lead, chromium, and cadmium are present in some plastic products as stabilizers and pigments; and these elements tend to end up concentrated in the ash. If the compounds of these metals in the ash are leachable, soil and groundwater contamination is possible. Advocates of incineration are convinced, however, that current technology will prevent stack discharge of most toxic and corrosive combustion products. They further claim that only a very small fraction of the heavy metals in the ash is leachable, and this should pose no problem if proper disposal procedures are followed. Moreover, heavy-metal-based pigments and stabilizers gradually are being phased out of many applications in favor of organic substitutes. Current trends suggest increasing reliance on the incineration approach despite the claimed drawbacks. There are approximately 135 waste incinerators in operation in the United States, processing about 13 percent of the nation's trash. It is estimated that by the year 2010 about half of all municipal solid waste will be incinerated in energy-recovery facilities.

Incineration is far more extensively practiced in Western Europe, with Germany handling over 40 percent of its trash this way. In Japan the figure is 68 percent.[101]

Recycling is the route under most intense development at this time because of its broad public appeal. It has obvious environmental advantages and is not opposed by any strong voting blocks or economic interests. In cases involving pre-consumer waste or scrap, where material identity and uniformity can be reasonably maintained, it often is possible to recycle the plastic back to the same product. In other instances, including those where carefully targeted post-consumer collection methods are possible, a secondary product can be made of a single recycled plastic. For example, PET beverage bottle scrap can be recycled to fabricate bottles for nonbeverage applications.[102] However, obtaining acceptable levels of segregation of individual plastic types from the waste recycle stream frequently is not practicable. In such instances, secondary products are made from commingled plastics waste. Commercial processes are in operation that convert commingled plastics, sometimes blended with virgin resins, fillers, and compatibilizers, into such varied products as lumber substitutes, waste baskets, pallets, computer diskette containers, and food trays.

Much of the future success of plastics waste recycling will depend on the development of effective collection and separation systems, which, along with appropriate incentives, will ensure the broad and willing participation of industry and consumers.

Another environmental issue, which also became of special concern during the 1980s, involves blowing agents used to make foamed plastics.[104] All blowing agents eventually escape to the atmosphere, and among them there is particular concern with the chlorofluorocarbons (CFCs), which have been linked to the depletion of the earth's stratospheric ozone layer. An international treaty signed in 1990 commits developed countries to phase out the manufacture of CFCs completely by 2000. In some more restricted geographical areas the smoke-forming potential of hydrocarbon blowing agents also is considered an issue.

Because CFCs have been widely employed in foamed plastics, including polystyrene, rigid and flexible polyurethanes, and polyolefins, for example, there has been intense activity to develop satisfactory substitutes. Among the most promising of these are the hydrochlorofluorocarbons (HCFCs), which have only 2 to 10 percent of the ozone depletion potential of CFCs, and hydrofluorocarbons, which have zero ozone depletion potential. The current consensus is that the HCFCs represent a transitional solution to the problem.[104] There has been promising development work with CO_2 as a blowing agent for polystyrene foam sheet.[105] In this application CO_2 is considered environmentally benign.

Along with the plastics waste issue, significant new developments can be anticipated as the industry continues its aggressive search for solutions to these important environmental problems.

REFERENCES

1. *Modern Plastics*, p. 113 (Jan. 1991).
2. *Facts and Figures of the U.S. Plastics Industry*, 1990 ed., p. 130, Society of Plastics Industry, Washington, DC, 1990.
3. Martino, R., *Modern Plastics*, p. 61 (Sept. 1990).
4. Ott, E. (Ed.), *Cellulose and Cellulose Derivatives*, p. 1093, Interscience, New York, 1943.
5. (a) Carothers, W. H., *Collected Papers*, H. Mark and G. S. Whitby (Eds.), John Wiley & Sons, New York, 1940.
   (b) Flory, P. J., *Principles of Polymer Chemistry*, Cornell University Press, Ithaca, NY, 1953.
6. Morawetz, H., *Polymers: The Origins and Growth of a Science*, John Wiley & Sons, New York, 1985.
7. (a) Kresta, J. E. (Ed.), *Polymer Additives*, Plenum Press, New York, 1984.
   (b) Scott, G., *Developments in Polymer Stabilization*, Applied Science Publishers, London, 1979.
8. Thayer, A. M., *Chem. Engineering News*, p. 37 (July 23, 1990).
9. Hawley, G. G. (Ed.), *The Condensed Chemical Dictionary*, 10th ed., p. 403, Van Nostrand Reinhold, New York, 1981.

10. International Union of Pure and Applied Chemistry, Macromolecular Division, Nomenclature Commission, *Pure Appl. Chem.*, **40**, 477 (1974); **48**, 374 (1976); **57**, 151 (1985).
11. Brandrup, J., in *Polymer Handbook*, 3rd ed., J. Brandrup and E. H. Immergut (Eds.), Vol. I, p. 1, John Wiley & Sons, New York, 1989.
12. Hill, L., in *Mechanical Properties of Coatings*, D. Brezinski and T. J. Miranda (Eds.), Federation of Societies For Coatings Technology, Philadelphia, PA, Apr. 1986.
13. Ward, I. M., *Mechanical Properties of Solid Polymers*, 2nd ed., John Wiley & Sons, New York, 1983.
14. (a) Mark, H., et al. (Eds.), *Encyclopedia of Polymer Science and Engineering*, 2nd ed., p. 101, John Wiley & Sons, New York, 1985.
    (b) ASTM-D-684-72.
15. Swallow, J. C., *The History of Polythene*, Renfrew and Morgan (Eds.), Iliffe and Sons, London, 1957.
16. *Paper Film and Foil Converter*, p. 9 (Aug. 1990).
17. *Modern Plastics*, p. 48 (July 1990).
18. Hogan, J. P., and Banks, R. L., U.S. Patent 2,825,721.
19. (a) Ziegler, K., and Gellert, H. G., U.S. Patents 2,699,457 and 2,695,327.
    (b) Ziegler, K., Holzkamp, E., Breil, H., and Martin, H., *Angew. Chem.*, **67**, 54 (1955).
    (c) Ziegler, K., and Martin, H., *Makromol. Chem.*, **18–19**, 186 (1956).
    (d) Natta, G., *J. Polym. Sci.*, **16**, 143 (1955).
    (e) Natta, G., Pino, P., Corradine, P., Danusso, F., Mantica, E., Mazzanti, G., and Moraglio, G., *J. Am. Chem. Soc.*, **77**, 1708 (1955).
20. Henrici-Olivé, G., and Olivé, S., *Chemtech*, p. 746 (Dec. 1981).
21. Albright, L. F., *Processes for Major Addition-Type Plastics and Their Monomers*, McGraw-Hill, New York, 1974.
22. Natta, G., and Danusso, G., (Eds.), *Stereoregular Polymers and Stereospecific Polymerization*, Pergamon Press, New York, 1967.
23. Natta, G., et al., *J. Am. Chem. Soc.*, **84**, 1488 (1962).
24. Leaversuch, R. D., *Modern Plastics*, p. 62 (Sept. 1990).
25. Rogers, J. K., *Modern Plastics*, p. 34 (July 1990).
26. (a) Ishihara, N., Kuramoto, M., and Loi, M., *Macromolecules*, **21**, 3356 (1988).
    (b) Zambelli, A., et al., *Macromolecules*, **22**, 104 (1989).
27. (a) Lindsay, K. F., *Modern Plastics*, p. 117 (Jan. 1990).
    (b) Illig, S. A., *Modern Plastics*, p. 112 (Jan. 1990).
28. Mark, H., et al. (Eds.), *Encyclopedia of Polymer Science and Engineering*, 2nd ed., Vol. 6, p. 123, John Wiley & Sons, New York, 1985.
29. *Plastics Technology*, p. 103 (Oct. 1990).
30. (a) Ranney, M. W., *Polyimide Manufacture*, Noyes Data Corp., Park Ridge, NJ, 1971.
    (b) Hergenrother, P. M., in *Step-Growth Polymerizations*; Solomon, D. H. (Ed.), Marcel Dekker, New York, 1972.
31. Farressey, W. H., Rose, J. S., and Carleton, P. S., *J. Appl. Polymer Sci.*, **14**, 1093 (1970).
32. *Modern Plastics*, p. 99 (Jan. 1990).
33. German Patent No. 728,981.
34. Backus, et al., in *Encyclopedia of Polymer Science and Engineering*, 2nd ed., H. Mark et al. (Eds.), Vol. 13, p. 243, John Wiley & Sons, New York, 1985.
35. Squiller, E. P., and Rosthauser, J. W., *Polymeric Materials Science and Engineering: Proceedings*, Vol. 55, p. 640, American Chemical Society, Washington, DC, 1986.
36. *Modern Plastics*, p. 21 (Oct. 1989).
37. Rogers, J. K., *Modern Plastics*, p. 50 (June 1990).
38. Wood, A. S., *Modern Plastics*, p. 55 (Oct. 1990).
39. Whiteside, R. C., Sheih, P. S., and Massingill, J. L., *J. Coating Technology*, **62**, 61 (1990).
40. Greenley, R. Z., in *Polymer Handbook*, 3rd ed., J. Brandrup and E. H. Immergut (Eds.), Vol. II, p. 153, John Wiley & Sons, New York, 1989.
41. Stevens, M. P., *Polymer Chemistry*, p. 151, Addison-Wesley, Reading, MA, 1975.
42. (a) Ivin, K. H., *Olefin Metathesis*, Academic Press, London, 1983.
    (b) Grubbs, R. H., in *Comprehensive Organometallic Chemistry*, p. 151, Addison-Wesley, Reading, MA, 1975.
43. (a) Klevins, H. B., *Chem. Revs.*, **47**, 1 (1950).
    (b) Harkins, W. D., *J. Chem. Phys.*, **13** (47), 381 (1945).
    (c) Klevins, H. B., *J. Am. Chem. Soc.*, **69**, 1928 (1947).
44. Simon, R. H. M., and Oster, B., U.S. Patent 3,345,430, 1967.
45. Mark, H., and Hohenstein, W. P., *J. Polymer Sci.*, **1**, 127 (1946).
46. Hardman, B., and Torkelson, A., in *Encyclopedia of Polymer Science and Engineering*, H. Mark et al. (Eds.), Vol. 15, pp. 204–308, John Wiley & Sons, New York, 1989.
47. Karol, F. J., presented at Plenary Session of ANTEC Meeting of the Society of Plastics Engineers, Dallas, TX, May 1990.

48. Bauer, E. A., and Collins, E. A., in *Rheology*, F. R. Eirich (Ed.), Vol. 4, p. 423, Academic Press, New York, 1967.
49. Lin, O. C. C., *J. Appl. Polym. Sci.*, **19**, 199 (1975).
50. Williams, M. L., et al., *J. Am. Chem. Soc.*, **77**, 3701 (1955).
51. Fox, T. G., et al., in *Rheology*, F. R. Eirich (Ed.), Vol. 1, p. 431, Academic Press, New York, 1956.
52. Cogswell, F. N., *Polymer Melt Rheology*, p. 40, Halsted Press, New York, 1981.
53. Carley, J. E., *Modern Plastics*, **39**, 123 (Dec. 1961).
54. Mendelson, R. A., in *Encyclopedia of Polymer Science and Technology*, N. M. Bikales (Ed.), Vol. 8, pp. 587–620, Wiley-Interscience, New York, 1968.
55. White, J. L., in *Rheometry, Industrial Applications*, K. Walters (Ed.), pp. 209–280, Research Studies Press, New York, 1980.
56. van Wazer, J. R., et al., *Viscosity and Flow Measurement, a Laboratory Handbook of Rheology*, Wiley-Interscience, New York, 1963.
57. Reiner, M., *Physics Today*, p. 62 (Jan. 1964).
58. Mooney, J., *J. Colloid Sci.*, **6**, 162 (1951).
59. Kruder, G. A., in *Encyclopedia of Polymer Science and Technology*, H. F. Mark et al. (Eds.), Vol. 6, pp. 571–631, John Wiley & Sons, New York, 1986.
60. Rauwendaal, C., *Polymer Extrusion*, Hanser, Munich, 1986.
61. Ibid., pp. 66, 246.
62. Ibid., p. 276.
63. Ibid., p. 456.
64. Ibid., pp. 352–441.
65. Colby, P. N., *Screw and Barrel Technology*, p. 22, Spirex Corporation, Youngstown, OH, 1985.
66. Michaeli, W., in *Encyclopedia of Polymer Science and Technology*, H. F. Mark (Eds.), Vol. 4, p. 828, John Wiley & Sons, New York, 1986.
67. Rubin, I. I., in *Encyclopedia of Polymer Science and Technology*, H. F. Mark et al. (Eds.), Vol. 8, p. 102, John Wiley & Sons, New York, 1987.
68. Ibid., p. 109.
69. Manzione, L. T., in *Encyclopedia of Polymer Science and Technology*, H. F. Mark et al. (Eds.), Vol. 14, pp. 86–90, John Wiley & Sons, New York, 1988.
70. Hedrick, R. M., et al., in *Reaction Injection Molding*, J. E. Kresta, (Ed.), ACS Symposium Series, No. 270, p. 135, American Chemical Society, Washington, DC, 1985.
71. Matejka, L., et al., *J. Applied Polymer Sci.*, **30**, 2787–2803 (1985).
72. Hull, J. L., in *Encyclopedia of Polymer Science and Technology*, H. F. Mark et al. (Eds.), Vol. 4, pp. 87–88, John Wiley & Sons, New York, 1986.
73. Meyer, R. W., *Handbook of Pultrusion Technology*, Chapman and Hall, New York, 1985.
74. Evans, D., in *Encyclopedia of Polymer Science and Technology*, H. F. Mark et al. (Eds.), Vol. 13, pp. 660–665, John Wiley & Sons, New York, 1986.
75. Irwin, C., in *Encyclopedia of Polymer Science and Technology*, H. F. Mark et al. (Eds.), Vol. 2, pp. 447–478, John Wiley & Sons, New York, 1985.
76. Rosato, D. V., in *Blow Molding Handbook*, D. V. Rosato and D. V. Rosato (Eds.), pp. 1–96, Hanser, Munich, 1988.
77. McConnell, W. K., Jr., *Principles and Practices of Thermoforming Technology*, McConnell Co., Ft. Worth, TX, 1987.
78. Throne, J. L., *Thermoforming*, Hanser, Munich, 1987.
79. McDonald, J. N., in *Encyclopedia of Polymer Science and Technology*, H. F. Mark et al. (Eds.), Vol. 16, pp. 810–821, John Wiley & Sons, New York, 1988.
80. Dodge, P. T., in *Encyclopedia of Polymer Science and Technology*, H. F. Mark et al. (Eds.), Vol. 14, pp. 659–670, John Wiley & Sons, New York, 1988.
81. Suh, K. W., and Webb, D. D., in *Encyclopedia of Polymer Science and Technology*, H. F. Mark et al. (Eds.), Vol. 3, p. 31, John Wiley & Sons, New York, 1985.
82. Ibid., pp. 28–29.
83. Ibid., pp. 25–26.
84. Ibid., pp. 38–41.
85. Meinecke, E. A., and Clark, R. C., *Mechanical Properties of Plastic Foams*, pp. 30–58, Technomic, Westport, CT, 1973.
86. French Pat. 1,498,620, Sept. 11, 1967, assigned to Ugine-Kuhlmann.
87. Gmitter, G. T., et al., in *Plastic Foams*, K. C. Frisch and J. H. Saunders (Eds.), pp. 153–176, Marcel Dekker, New York, 1972.
88. Benning, C. J., *Plastic Foams*, pp. 126–138, Wiley-Interscience, New York, 1969.
89. Ibid., pp. 2–116.
90. Ingram, A. R., and Fogel, J., in *Plastic Foams*, K. C. Frisch and J. H. Saunders (Eds.), pp. 531–597, Marcel Dekker, New York, 1972.

91. Suh, K. W., U.S. Patents 3,897,528 (July 29, 1975) and 3,914,085 (Oct. 21, 1975), assigned to Dow Chemical Co.
92. Tadmor, Z., and Gogos, C. G., *Principles of Polymer Processing*, pp. 351–361, 467–507, John Wiley & Sons, New York, 1979.
93. Carley, J. F., and Stroub, R. F., *Ind. Eng. Chem.*, **45**, 970–973 (1953).
94. Carley, J. F., et al., *Ind. Eng. Chem.*, **45**, 974–977 (1953).
95. Martelli, F. G., *Twin-Screw Extruders: A Basic Understanding*, Van Nostrand Reinhold, New York, 1983.
96. Callari, J. J., *Plastics World*, pp. 12–18 (Sept. 1989).
97. Klemchuk, P. P., *Polymer Degradation and Stability*, **27**, 183–202 (1990).
98. Francis, S., *Plastics Compounding*, pp. 20–38 (Nov./Dec. 1989).
99. Thayer, A. M., *Chem. & Eng. News*, pp. 7–14 (June 25, 1990).
100. Nir, M. M., *Plastics Engineering*, pp. 21–28 (Oct. 1990).
101. Erickson, D., *Plastics World*, pp. 39–43 (Sept. 1989).
102. *Modern Plastics*, pp. 45–47 (Jan. 1991).
103. Voss, D., *Chemical Engineering Progress*, pp. 67–72 (Oct. 1989).
104. Golanty, D. *Plastics Compounding*, pp. 24–33 (July/Aug. 1991).
105. Welsh, G. C., European Patent Application 0 411 923 A2 (to Dow Chemical Co.), Aug. 1, 1990.

SELECT BIBLIOGRAPHY

Biesenberger, J. A., and Sebastian, D. H., *Principles of Polymerization Engineering*, John Wiley & Sons, New York, 1983.
Blow Molding Handbook, D. V. Rosato and D. V. Rosato (Eds.), Hanser, Munich, 1989.
Encyclopedia of Polymer Science and Engineering, 2nd ed., H. Mark et al. (Eds.), John Wiley & Sons, New York, 1985.
Facts and Figures of the U.S. Plastics Industry, 1990 ed., Society of the Plastics Industry, Washington, DC, 1990.
Kirk-Othmer Encyclopedia of Chemical Technology, 3rd ed., H. Mark et al. (Eds.), John Wiley & Sons, New York, 1978.
Michaeli, W., *Extrusion Dies*, Hanser, Munich, 1984.
Odian, G., *Principles of Polymerization*, McGraw-Hill, New York, 1970.
Plastics Foams, Vols. 1 and 2, K. C. Frisch and J. H. Saunders (Eds.), Marcel Dekker, New York, 1972.
Plastics, Thermoplastics and Thermosets: A Desk-Top Data Bank, 9th ed., International Plastics Selector, San Diego, CA, 1987.
Polymer Handbook, 3rd ed., J. Brandrup and E. H. Immergut (Eds.), John Wiley & Sons, New York, 1989.
Rauwendaal, C., *Polymer Extrusion*, Hanser, Munich, 1986.
Reaction Engineering of Step Growth Polymerization, S. K. Gupta and A. Kumar (Eds.), Plenum, New York, 1987.
Seymour, R. B., and Carraher, C. E., Jr., *Polymer Chemistry*, 2nd ed., Marcel Dekker, New York, 1988.
Tadmor, Z., and Gogos, C. G., *Principles of Polymer Processing*, John Wiley & Sons, New York, 1979.
Ullmann's Encyclopedia of Industrial Chemistry, 5th ed., W. Gerhartz (Ed.), UCH Publishers, Deerfield Beach, FL, 1985.

20

The Chemistry of Structural Adhesives

Denis J. Zalucha*

INTRODUCTION

Adhesives have been used successfully in a variety of applications for centuries. Today, adhesives are more important than ever in our daily lives, and their usefulness is increasing rapidly. In the past few decades there have been significant advances in materials and in bonding technology. People now routinely trust their fortunes and their lives to adhesively bonded structures and rarely think about it, but the subject of adhesives and adhesive bonding is of great importance to researchers.

The scientific literature on adhesives and adhesion is growing at a rapid pace. There are many texts devoted to adhesion and adhesives.[1-11] New patents on adhesive compositions and processes are granted almost daily, and there are scientific journals devoted to the science and technology of adhesives and their use. A significant portion of the publications dealing with adhesives are concerned with epoxy, urethane, and acrylic structural adhesives, as they are used in a wide variety of commercially important applications. This chapter reviews some of the chemistry of these adhesive types.

Adhesion

Materials are generally defined as adhesives by what they do. Almost any organic polymer and even many inorganic materials can function as adhesives in some situations. However, whether they are organic polymers or inorganic, or perhaps even solders or brazing alloys, all adhesives share common traits in performing their functions:

1. An adhesive, by surface attachment only, transfers and distributes mechanical loads among the components of an assembly.
2. At some time in the course of the bond formation, the adhesive must be liquid or behave as a liquid in order to wet the adherends.
3. An adhesive carries some continuous, low, and often variable load throughout its life.
4. An adhesive must work with the other components of the assembly to provide a durable product that is resistant to degra-

*S. C. Johnson Polymer, Racine, WI.

dation by elements of the environment in which it will be used.

We all seem to have some built-in operational definition of "adhesion." If we are given a bonded assembly to test or evaluate, we usually bend, twist, and stretch it to try to make it fail. We typically equate adhesion with resistance to failure: the greater the resistance, the better the adhesion, and, by implication, the better the adhesive. This is not a bad working definition of adhesion and serves adequately in most cases. However, there are instances in which one wants assemblies to come apart fairly easily, such as protective masking or packaging. In these cases one might regard too much resistance to destruction as a mark against the adhesive.

The expectations of the user are extremely important in determining whether an adhesive is "good" or "bad." Adhesives are really judged on the ability of the whole assembly to meet the user's expectations, which will, in turn, depend on the way the assembly is loaded and tested and on what and where are the weakest points of the assembly. Adhesion is not an intrinsic property of any polymer but is rather a property of the whole assembly. In a broad sense, adhesion is really a number we measure or assign to the response of the assembly to some destructive deformation. Thus, adhesion is defined by the test, and the result of the test (the number we assign) is determined by the weak points or flaws in the structure. The acceptability of the bonded assembly is determined by our expectations of how the adhesive will respond to the destructive forces.

Curing

The chemistry of a structural adhesive is designed to do at least two important things. First, the adhesive must at some time pass through a fluid state in order to wet the adherends. Second, the adhesive in its final state in the bond line must be a solid, high molecular weight polymer that is able to carry and transfer mechanical forces. The chemistry must make possible some manufacturing process for the assembly that allows for the liquid state, and there must be some mechanism for passing from the liquid to the solid, load-bearing state. The process of going from the liquid to the solid state usually is termed "curing."

There are three general ways in which adhesives are cured. The adhesive can be applied in a molten state and allowed to cool and solidify in the bond line. In a second method, the adhesive can be applied as a solution or dispersion and the carrier liquid allowed to evaporate, leaving behind the high molecular weight polymer. In the third method the adhesive consists of a low viscosity fluid containing reactive groups that undergo polymerization in the bond line to build the molecular weight sufficiently to carry a load.

The third method is the one most commonly used with epoxy, urethane, and acrylic structural adhesives. The chemistry, methods and mechanisms for accomplishing this polymerization will be covered in more detail in the following sections.

Adhesion Mechanisms

Once the liquid adhesive has been applied to the adherends and intimate contact and wetting have been established, the liquid mass is cured by polymerizing it to a solid, high molecular weight, load-bearing state. It now transfers load among the components of the assembly. But we are interested here in what actually accounts for the adhesion between the polymer and the adherend surface, the ability to transfer load. Three mechanisms have been postulated to explain the adhesion of one material to another: (1) mechanical interlocking, (2) electrostatic attraction, and (3) the formation of chemical bonds across the interface. It is probable that all three mechahisms play some role in any bonding process although often one contributes much more than the others.

Mechanical interlocking probably plays a part in the adhesion of ice to glass. Silver amalgam dental fillings are held in place largely by mechanical interlocking. Although there probably are some exceptions, mechanical

interlocking usually is not a major factor in bond formation with structural adhesives.

The other extreme in bonding is the formation of direct chemical links across the interface. These bonds would be expected to be quite strong and durable, but they are difficult to form. It is necessary that there be mutually reactive chemical groups tightly bound on the adherend surface and in the adhesive, and there is evidence that such bonds can be formed under controlled conditions. Silane coupling agents are one example of using specific reactive groups to promote the formation of direct chemical bonds.

By far the dominant adhesion mechanism, particularly in the absence of tightly bound reactive groups, is the electrostatic attraction of the polar groups of the adhesive to polar groups of the adherends and to each other. These are mainly transient dispersion forces (London forces) and forces arising from the interaction of permanent dipoles. These forces provide much of the attraction between the adhesive and the adherend and also provide a significant portion of the cohesive strength of the adhesive polymer. Dispersion forces arise from the interaction of trasient dipoles present in all atoms and molecules, whether permanently charged or not. Dispersion forces act over a very short range and decrease with the sixth or seventh power of distance. The attractive forces between permanent dipoles can be somewhat stronger but still decrease with the sixth power of distance. With such a great dependence on distance, neither force is significant beyond about 0.5 nanometer. The need to obtain intimate contact between the adhesive and the adherend surface is obvious.

In two articles on the cohesive and adhesive strengths of polymers, H. F. Mark[12,13] derives some estimates of what adhesive bond strengths might be achievable with covalent bonds or polar forces across an interface. He concludes that the bonds actually achieved in real life are only a small fraction of what he estimates for the situation in which covalent chemical bonds are the main contributor to adhesion. He further proposes that even if there are a significant number of covalent chemical bonds across the interface, the failing strength of the bond still will depend on the strength of the polar bonds. The polar bonds will fail at a lower strength than the covalent bonds, and the applied load then will be concentrated on the covalent bonds. The measurable mechanical strength of a partially covalent adhesive bond still will be dominated by the polar forces. The implication is that although increasing the proportion of covalent bonds across the interface can lead to enhanced durability, the ultimate load-carrying capacity probably will not be significantly affected.

Surfaces

Adhesives must function solely through surface attachment. Therefore, the nature of the condition of the adherend surface are crucial to the formation of strong and durable bonds. By "surface" we usually mean that region of a material which interacts with its surroundings. There is some region of a bonded assembly where the adhesive and the adherend interact, but only rarely is this a sharp boundary. Usually it is a very diffuse, somewhat ill-defined region of interaction that has become known as an interphase rather than an interface. For example, the interaction with a freshly cleaved single crystal of zinc might occur over only a few atomic layers or a few nanometers. Rough or porous surfaces present more surface area than smooth ones of the same dimensions, and the adhesive might reach a depth of several hundred nanometers on a porous adherend such as wood or paper.

One very important aspect of surfaces is that they rarely have the same chemical composition as the bulk material and often seem to be entirely unrelated to the bulk. The surface usually consists of several regions having no clear boundaries. A metal alloy might have a well-defined bulk composition, but at the surface there probably will be a region that is still metallic but is of different chemical composition because of alloying elements or impurities that have segregated at the surface. On top of this region there probably will be a layer of oxides and hydroxides formed by reaction with the

atmosphere. There also will be many other contaminants such as nitrogen, sulfur, and halogen compounds formed by interaction with the pollutants in the atmosphere. Finally there will be several layers of adsorbed water. The surface of a metal also might be contaminated with rolling oils, cutting lubricants, drawing compounds, or corrosion inhibitors. Mechanical working of the metal might even mix these contaminants with the other surface materials to create something like a nonhomogeneous "frosting" on the surface.

Engineering plastics display some of the same surface phenomena as metals, in that the surface is very different from the bulk. The manufacturing process often introduces anisotropy so that the mechanical properties of the material are different in different directions. In addition, it is common to find that components of the plastic have accumulated at the surface. Low molecular weight polymers or oligomers, plasticizers, pigments, mold release agents, shrink control agents, and other processing aids as well as adsorbed contaminants often are present.

The nature of the surface of an engineering plastic can change rapidly in response to its surroundings. The bulk of the material might be in the glassy state, but because of the concentration of low molecular weight material and contaminants, the surface region can be quite mobile. Exposure of the surface to a polar environment, such as by wiping with a polar solvent, can cause polar groups in the plastic surface to preferentially orient themselves outward. Exposure to a nonpolar medium can bring out the nonpolar nature of the surface.

The adhesive layer itself can be thought of as one example of an engineering plastic. There is no reason to expect that the adhesive should be homogeneous throughout. Especially because the adhesive must go through a liquid, highly mobile state and must intimately interact with the adherend surface, it is quite likely that the adhesive in the interphase region is quite different from the bulk adhesive. Low molecular weight materials can be drawn into the surface of a porous adherend, leaving higher-weight polymer and fillers behind. The polar or nonpolar nature of the adherend influences the orientation and morphology of adhesive polymers in the interphase. Compounds at the adherend surface can catalyze or inhibit polymerization. Solvents in the adhesive can swell the adherend or dissolve portions of the adherend surface.

The interphase region is complex, and its composition usually is unknown. Primers or surface treatments often are used to improve control of the interphase and provide increased adhesion, durability, and resistance to aggressive environments. The chemistry of primers and surface treatments is as varied as the chemistry of the adhesives.

Any bonded construction consists of at least two adherends and one adhesive and contains at least two interphase regions. It is important to remember that the performance of the construction, its durability, its mechanical properties, and its response to tests and challenges, are all properties of the entire assembly. The successful use of adhesives depends on taking account of all parts of the construction and the process. Whereas the adhesive is just one part of the assembly, its chemistry plays an important role in the bonding process.

The following sections discuss the chemistry of some major classes of adhesives.

EPOXY STRUCTURAL ADHESIVES

Introduction

Epoxy adhesives command a large portion of the structural adhesives market. Most people probably are familiar with epoxy structural adhesives, as these are the typical two-part adhesives found in hardware stores and supermarkets. Epoxy adhesives owe their popularity with both the general public and industry to their ease of use, their relative safety, and their compatibility with many adherends. The various chemical reactions involving the epoxy ring provide a fertile field for the development of a wide range of properties.[14] Even a half century after the first epoxy

patents were issued,[15] new patents on epoxy adhesive technology continue to appear every month.

Epoxy adhesives get their name from the 1,2-epoxy, epoxide, or oxirane ring, which provides the reactive sites on the polymers. This three-member ring consists of two carbon atoms joined to an oxygen atom. The strained geometry contributes to the reactivity with nearly all nucleophilic compounds, not only curatives but also adherend surfaces.

Epoxy ring

Typical epoxy resins used to formulate epoxy adhesives have at least two epoxy rings, usually at the ends of a relatively short-chain prepolymer. The epoxy groups then are reacted with other epoxy groups in a chain-growth polymerization or with another curative in a step-growth polymerization to produce a polymer network, which can be either thermoplastic or thermoset. The polymer linkages created by reaction of the epoxy ring are polar and provide adhesion to a variety of polar surfaces. With the proper backbone polymers and curing agents, cured epoxy adhesives can be very tough and resistant to chemical degradation. Bonded assemblies can have a high degree of durability and environmental resistance. Figure 20.1 shows the hand application of an epoxy adhesive to an automotive sheet molding compound (SMC) hood assembly.

Commercial Synthesis of Epoxy Resins

Epichlorohydrin, or 1-chloro-2,3-epoxypropane, is a key raw material in the manufacture of epoxy resins. Epichlorohydrin is made by chlorination of propylene, treatment with hypochlorous acid, and dehydrohalogenation of the intermediate to give the mono-epoxide.

Epichloroydrin

Most epoxy resins used in adhesives are made by reacting a compound containing an active hydrogen atom with epichlorohydrin.

Fig. 20.1. Hand application of a two-part epoxy adhesive to a SMC (sheet molding compound) automotive hood assembly. (*Courtesy Lord Corp., Erie PA.*)

This intermediate then can be dehydrohalogenated to give an epoxy functional material. Usually the starting material will contain at least two reactive hydrogen atoms to give a product having at least two epoxy groups per molecule.

Epoxides can also be prepared by direct chlorination of an olefin followed by dehydrohalogenation with base, or by the direct oxidation of an olefin by a peracid such as peracetic acid prepared *in situ* from the acid and hydrogen peroxide.

The most commercially important epoxy resins for adhesives are the diglycidyl ether of bisphenol A (DGEBA) and its higher homologues. Bisphenol A is made by condensing phenol and acetone with an acid catalyst. The bisphenol A then is reacted with epichlorohydrin, with dehydrohalogenation with sodium hydroxide following to give the product.

Diglycidyl ether of bisphenol A

The NaOH is present only in catalytic amounts until the addition of the epichlorohydrin is complete; then it is increased to a stoichiometric amount for the dehydrohalogenation. If this reaction is carried out with a large excess of epichlorohydrin, the principal product is the simplest diglycidyl ether. The pure DGEBA has a molecular weight of 340 and an epoxide equivalent weight of 170, and it melts at about 42°C. Because there are some side reactions, typical commercial resins have molecular weights of about 370 and epoxide equivalent weights in the range of about 180 to 190, and will have some small percentage of material that is less than difunctional. The diglycidyl ether of bisphenol A is sold by several companies, such as Celanese (Epi-Rez 510), Reichhold (Epotuf 37-140), Shell (Epon 828), CIBA-GEIGY (Araldite GY 6010), and Dow Chemical (D.E.R. 331).

If the excess of epichlorohydrin is decreased and the reaction run with nearer the stoichiometric amounts of reactants, the molecular weight of the product increases. The higher molecular weight products are homopolymers of DGEBA, having pendant secondary hydroxyl groups. They can be produced commercially either through the direct reaction of bisphenol A, NaOH, and epichlorohydrin or by chain extension of the DGEBA with bisphenol A. The value of n can be as high as 30 for some resins. The higher-weight polymers are generally not practical in adhesives because the number of epoxy groups available for cross-linking becomes small, and the higher glass transition temperatures and lower reactivities make it impractical to drive curing reactions to completion.

Bisphenol A resin

Another important class of aromatic epoxy resins is the epoxy novolac resins. Novolacs are polymers made by condensing phenol or cresol with formaldehyde in the presence of an acid catalyst. The result is a chain of phenol or cresol groups held together by methylene bridges in a mixture of *ortho* and *para* isomers. The hydroxyl groups then can be reacted with epichlorohydrin and dehydrohalogenated with base to produce an epoxy functional novolac. The simplest, containing only two rings, is the diglycidyl ether of bisphenol F. It is a liquid resin having an epoxy equivalent weight of about 165.

Value of n for epoxy novolac resins used in adhesives are usually in the range of about 0.2 to about 3.5.

Low-viscosity monofunctional reactive diluents also are important components of many formulated epoxy structural adhesives. They are used to lower the viscosity of the composition without introducing low molecular weight unreactive species that might migrate out of the adhesive after it is cured. These compounds are made in much the same way that the polyfunctional epoxy resins are made. Some of the common reactive diluents are butyl glycidyl ether, cresyl glycidyl ether, and alkyl glycidyl ethers, among others.

Diglycidyl ether of bisphenol F

Unlike that of the higher molecular weight bisphenol A resins, the epoxy functionality of the epoxy novolac resins remains relatively constant with increasing molecular weight because all of the aromatic rings contain hydroxyl groups that can be epoxidized. Higher molecular weight epoxy novolac resins can produce adhesives with higher cross-link densities.

Butyl glycidyl ether

Cresyl glycidyl ether

Epoxy phenol novolac

Alkyl glycidyl ether

Butyl glycidyl ether has the lowest viscosity and gives the greatest viscosity reduction at the least concentration. However, it has a higher vapor pressure than either cresyl glycidyl ether or the higher alkyl glycidyl ethers, and is more likely to cause problems with skin sensitivity and toxicity. For safety reasons the higher molecular weight, lower vapor pressure materals find more use even though they might be less efficient than those with a higher vapor pressure.

The diglycidyl ether of 1,4 butanediol is a commercially available (CIBA-GEIGY's RD-2) low-viscosity difunctional epoxy that is a somewhat less efficient diluent than the mono-epoxides but does offer the advantage of having two reactive sites.

Epoxidized oils such as soybean oil or linseed oil, which are made by oxidizing the internal unsaturation in the fatty acid chain, find much use as plasticizers, particularly in vinyl resins. An internal epoxy ring is much less reactive than a terminal epoxy ring and usually does not take part in typical epoxy curing reactions under mild conditions. These materials can be used as plasticizers in epoxy adhesives but cannot really be considered reactive diluents.

Aliphatic epoxy compounds find some limited use in epoxy structural adhesives. They generally are low-viscosity liquids, so they can reduce viscosity with less impact on cross-link density than the monofunctional dliuents have. The cycloaliphatic compounds such as 3′,4′-epoxycyclohexylmethyl 3,4-epoxycyclohexane carboxylate (CY-179 from CIBA-GEIGY) are made by direct epoxidation of the corresponding cyclohexane with peracetic acid. They produce high T_g polymers too brittle for use as the sole epoxy compound in adhesives.

3′,4′-Epoxycyclohexylmethyl 3,4-epoxycyclohexane carboxylate

Tougher polymers of somewhat lower T_g are obtained from glycidyl esters such as the diglycidyl ester of hexahydrophthalic acid (CY-184 from CIBA-GEIGY). The glycidyl esters are prepared by reacting the corresponding cycloaliphatic carboxylic acid with epichlorohydrin and dehydrohalogenating with NaOH.

Diglycidyl ester of hexahydrophthalic acid

There are also a variety of other specialty epoxy resins used in limited quantity for applications demanding some specific performance property. Tri- and tetra-functional materials such as triglycidyl-*p*-amino phenol, triglycidyl isocyanurate, or *N,N,N′,N′*-tetraglycidyl-4,4′-diaminophenyl methane (tetraglycidyl methanedianiline) have been used in adhesives requiring high heat resistance and good chemical resistance.

Tetraglycidyl methanedianiline

A typical sales specification for an epoxy resin used in adhesives will include the epoxy equivalent weight (EEW), also sometimes called the weight per epoxy (WPE), viscosity, and density at some specified temperature and the average functionality or number of epoxy groups per molecule. Sometimes a specification on total chlorine is included, which gives some indication of bound chlorine not removed by the NaOH in the dehydrohalogenation process. This bound chlorine is the result of epichlorohydrin side reactions during the epoxy synthesis. A high chlorine content indicates that there will be a large number of molecules of low functionality because each chlorine remaining represents one epoxy ring not formed in the dehydrohalogenation step.

Epoxy Cure Chemistry

Epoxy structural adhesives rely on the chemical reactions of the epoxy group with other reactants to pass from a liquid, wetting state to a solid, load-bearing state. There are a number of ways in which this is done, but all fall under one, or some combination, of three general schemes: step-growth polymerizations through reaction with curing agent, chain-growth polymerization initiated by Lewis acids, or chain-growth polymerizations initiated by Lewis bases.* Often the cure times of the slower step-growth curing adhesives are shortened by including Lewis acid or Lewis base catalysts.

It would be tempting to consider the step-growth and chain-growth polymerization reactions as if they were independent and one could have the choice of either in any particular situation. The truth is that there are aspects of both types of polymerization in the cure of almost every epoxy structural adhesive. Such multiple-cure reactions often make it difficult to calculate the stoichiometry of an epoxy adhesive formulation. One type might predominate, depending on the formulation and cure conditions, but the effects of the other could not be completely discounted. The significance of this statement can be seen by looking at the two generalized reactions.

A typical step-growth reaction is one in which a single epoxy ring reacts with the active hydrogen of the curing agent. The general reaction is:

Model epoxy reaction

In order to form a high molecular weight polymer, both the epoxy-containing material and the curative must be at least difunctional. If both behave as difunctional materials, the resulting polymer is linear, and then it is necessary to drive the reaction nearly to completion to obtain a high-enough molecular weight to be useful. Neither of the reactants can contain a significant amount of monofunctional material. Monofunctional impurities or reactive diluents will act as chain terminators and limit the ultimate molecular weight obtainable.

Chain-growth polymerizations occur through the reaction of epoxy rings with the active site on a growing chain and not with each other or with a second curing agent. After being initiated by a Lewis acid or a Lewis base, the growing chain will continue to consume epoxy groups and can reach a high molecular weight very rapidly. Because the epoxy groups are reacting with the growing chain, even a monofunctional epoxy compound can be polymerized. If A* is the initiator and M the monomer unit, the general reaction is:

$$A^* + nM \rightarrow A - M_{n-1}M^*$$

Chain-growth reaction

where the asterisk indicates the active site of the growing chain.

Lewis bases initiate anionic chain growth polymerizations, the generalized reaction being that of a propagating alkoxide anion.

Anionic chain growth

*Lewis acids and Lewis bases are discussed below in the section on "Chain-Growth Polymerizations."

Lewis acids initiate cationic chain-growth polymerizations. There are several possible chain propagation reactions, and the mechanism of cationic chain growth is still open to some debate. However, the propagating species is a cation formed by reaction of the Lewis acid initiator with an epoxy ring.

In chain-growth polymerizations, epoxy reactants containing more than one epoxy ring per molecule can form tightly cross-linked, three-dimensional networks, as the epoxy groups on one molecule can easily react with more than one growing chain.

Step-Growth Polymerization

Only a relative few of the dozens of active hydrogen compounds that undergo reactions with the epoxy ring find widespread use in epoxy structural adhesives. The most common are amines, amides, acid anhydrides, phenols, thiols, and carboxylic acids.

Primary and secondary amines react with epoxy groups to form secondary or tertiary amine linkages. The reactions proceed at a reasonable rate at room temperature with the glycidyl ethers and glycidyl esters. The resulting polymer depends on the structure of the reactants and the degree of cure, and thus a variety of final adhesive properties can be had. Because the reaction is catalyzed by hydrogen ions, weakly acidic compounds such as phenol often are used to accelerate the reaction. The reaction of the secondary amine with the epoxy group produces a tertiary amine, which can in turn be the Lewis base that initiates an anionic chain growth polymerization of the remaining epoxy groups, depending on cure conditions.

Epoxy amine reaction

The simple linear aliphatic diamines, H_2N-$(CH_2)_nNH_2$, can be used as curatives in adhesives. For small values of n, the short distance between the amine group can hinder the reaction of the second amine and slow the cure process. Also, the resulting products tend to be brittle for values of n less than about 6.

More flexible, tougher products can be obtained by using liquid diamines or polyamines having more flexible backbones. For example, Jeffamine T.403 (Texaco Chemical) is a low-viscosity liquid (70 mPa · s) having a molecular weight of about 440 and an amine equivalent weight of about 145. It is a poly(propylene oxide) triamine made from the polyether initiated by trimethylol propane, and is promoted as a flexible cross-linker for epoxy systems.

Various polyamines also are useful as curatives for epoxy adhesives. An example of a simple polyamine is diethylenetriamine (DETA):

$$H_2NC_2H_4NHC_2H_4NH_2$$

Diethylenetriamine

Its higher-weight homologues also are quite useful and are made by adding (CH_2CH_2NH) groups, leading to a homologous series of the form:

$$H_2N\text{-}(C_2H_4NH)_n\text{-}C_2H_4NH_2$$

Where n is two, the material is triethylenetetraamine (TETA), and for $n = 3$ it is tetraethylenepentamine (TEPA), and so on. These three are the most important members of this series for adhesive applications, and are used more often than the simple aliphatic diamines. They often are used in combination with some other curative. Calculating stoichiometry can be difficult with these polyamines. All the active hydrogens might not be available because of steric factors introduced by the first reactions, or because once the first one or two active hydrogens are reacted with epoxy rings, the molecule is anchored into the chain and cannot readily diffuse to other epoxy groups. Therefore, an excess of amine often is added to the adhesive formulation. (Too much excess amine can lead to soft adhesives because all the epoxy rings are used up before the molecular weight becomes very high.)

Other series of polyamines can be made too. For example, some or all of the ethyl groups can be replaced by propyl groups, or the compound might be modified by reaction with

an excess of a monoepoxide to give a hydroxy functional amine. Such modifications are made to improve adhesive properties and sometimes to lower toxicity of the curing agent or make it easier to handle.

Among the most important amine functional epoxy curing agents are the polyamide resins. These are made from dimerized unsaturated fatty acids by reaction of the dimer acid with a polyamine such as diethylenetriamine.

A variety of aromatic polyfunctional amines also are used in curing epoxy adhesives. They generally are slower to react than the aliphatic amines and require a heat cure to be practical. They do provide generally better high-temperature properties than the aliphatic amines. The most commonly used aromatic amine curatives are *m*-phenylenediamine (MPDA), methylenedianiline (MDA), and diaminodiphenyl sulfone (DDS).

$$HOOCC_{34}COOH + H_2NC_2H_4NHC_2H_4NH_2 \longrightarrow$$

$$H_2NC_2H_4NHC_2H_4NH\left(\overset{O}{\overset{\|}{C}}C_{34}\overset{O}{\overset{\|}{C}}NHC_2H_4NHC_2H_4NH\right)_n H$$

Polyamide Polyamine synthesis

These polyamides are available from several suppliers worldwide and are among the most common curatives in the general purpose, "do-it-yourself" two-package epoxy adhesives. They have a distinctive odor somewhat like popcorn and are easily recognized in adhesive formulations. Chemically they are polyamides, but the activity with epoxy resins is primarily through the amine groups and not the amide hydrogens. The polyamide backbone does contribute to the overall good mechanical properties of the polyamide amine cured adhesives.

The polyamide amines are very high-viscosity liquids, some having viscosities over 50,000 mPa · s. Typical amine equivalent weights are 100 to 150. The polyamide amines react with bisphenol A epoxy resins at room temperature although the adhesives usually require several hours to reach sufficient molecular weight to carry a load. Cure times can be shortened to a few minutes at about 150°C. Tertiary amine catalysts such as dimethylaminomethylphenol also can be used to shorten cure time.

Epoxy adhesives cured with polyamide amines are flexible, tough, durable adhesives useful on a wide variety of adherends. They probably have contributed heavily to making the words "epoxy" and "adhesive" equivalent for many people.

m-Phenylenediamine

Methylenedianiline

Diaminodiphenyl sulfone

All of the curatives described so far are used to make two-package adhesives in which the curing agent is packaged separately from the epoxy resin. Once they are mixed, they have a limited pot life, usually less than a few hours. It is possible to make one-package epoxy adhesives that can have very long shelf lives at room temperature but cure rapidly when heated. One amine curative widely used to make single-package heat-cured epoxy adhesives is dicyandiamide (cyanoguanadine), commonly known as dicy:

$$H_2N\overset{NH}{\overset{\|}{C}}NHC{\equiv}N$$

Dicyandiamide

It is made by dimerizing cyanamide in basic aqueous solution, and is a colorless solid

melting at 208°C. Dicyandiamide is soluble in polar solvents, but at room temperature is insoluble in bisphenol A epoxy resins. It can be made into a very fine powder and milled into epoxy resins to form stable dispersions. Because the dicy is insoluble in the epoxy, the only possible reaction sites are at the particle surfaces. Although some reaction certainly occurs over a short time, the adhesives easily can have a useful shelf life of six months. On heating to about 150°C, the dicyandiamide becomes soluble in the epoxy resin, and the adhesive polymerizes rapidly.

Carboxylic acids can be used to cure epoxy adhesives or otherwise modify epoxy adhesives. The reactions can be complex. If no hydroxyl groups are present initially, the first reaction will be that of the active hydrogen with the epoxy ring to form an ester. This will produce a hydroxyl group on the backbone and allows for competing reactions. The organic acid can catalyze the etherification reaction with the hydroxyl group or undergo a condensation esterification reaction directly with the hydroxyl group.

Epoxy acid reaction

Etherification reaction

Condensation reaction

These reactions usually are slow at room temperature, and the adhesives must be cured with heat.

Acid anhydrides also can be used to cure epoxy adhesives although they usually are used only where good service at high temperatures is required. Most of the anhydride-cured epoxy adhesives are cured at high temperature. Because most of the anhydrides are relatively small molecules, the products tend to be tightly cross-linked and can be somewhat brittle.

The first step in the anhydride cure is ring opening of the anhydride by an active hydrogen, perhaps from water or hydroxyl groups already present on the epoxy resin. The resulting acid then reacts as a typical organic acid.

Anhydride ring opening by hydroxyl

An interesting anhydride-cured composition was reported,[16] where 3,3′,4,4′-benzophenone tetracarboxylic dianhydride (BDTA) was found to form a liquid solution with triglycidyl *p*-aminophenol (TGpAP) that cured at room temperature to give a product with good high-temperature properties.

Triglycidyl *p*-aminophenol

3,3′,4,4′-Benzophenone tetracarboxylic dianhydride

Another group of popular curing agents for epoxy resins is the thiol-terminated polysulfides. Thiols react with glycidyl ethers quite rapidly, particularly when a catalyst such as tris-(dimethylaminomethyl) phenol is present.

The reaction is similar to that found with hydroxyl groups and produces a polythioether product. These are the familiar "five-minute" epoxy adhesives and have the characteristic odor of thiols.

Chain-Growth Polymerizations

Chain-growth polymerizations are very important to many commercially successful epoxy structural adhesives. They can be extremely rapid and contribute to the fast cure times needed for high productivity in many manufacturing operations.

A Lewis base is a compound that contains an unshared pair of electrons capable of undergoing chemical reactions. Tertiary amines are examples of Lewis bases, and often are used in epoxy curing agents. In an anionic epoxy polymerization the propagating species is the alkoxide anion generated by the reaction of the Lewis base with an epoxy ring.

Formation of anionic chain initiator

This anion can continue to react with epoxy rings, adding them to the chain until the anion is destroyed in some side reaction.

Anionic chain propagation

The product is a polyether, which can be tightly cross-linked when polyfunctional epoxides are used, as each epoxy ring can become part of a different chain.

Lewis acids, compounds with empty orbitals capable of accepting electron pairs, initiate cationic polymerization of epoxy resins. In this case the propagating species is a positive ion. The most commonly used Lewis acids are the boron trihalides, particularly BF_3 and BCl_3. They usually are used in the form of salts because both are gases at room temperature and are so reactive with epoxy resins that they can be difficult to control. Lewis base salts with the boron trihalides have much lower reactivity at room temperature but can react quickly on relatively mild heating, depending on the particular salt. Boron trifluoride readily forms salts with ethers, alcohols, and amines, and several of these salts are commercially available. A boron trifluoride ethanolamine salt can be included in the curative portion of a two-package epoxy adhesive. When the two packages are mixed at the time of use, the mixture can have a pot life of hours at room temperature but polymerize in minutes at temperatures of 100 to 150°C.

One very interesting new application of Lewis acids in curing epoxy adhesives has appeared within the last 15 years. The Lewis acid initiator for the cationic polymerization is formed by the heat or ultraviolet light-induced decomposition of Lewis acid:Lewis base salts. Several patents by Crivello and coworkers[17–21] describe compounds containing aromatic onium salts such as iodonium or sulfonium, in which the cation is stabilized by the aromatic rings.

Diphenyl iodonium salt

The counterion generally is a large stable anion such a $[SbF_6]^-$, $[AsF_6]^-$, $[PF_6]^-$, or, $[BF_4]^-$. The more stable anions are less likely to terminate the growing cationic chains than are typical anions. The salts can be dissociated with heat to release the cation, which appears capable of initiating cationic polymerization of many materials in addition to epoxy rings. If the proper dye sensitizers are added, the cation can be liberated by ultraviolet light to initiate the polymerization. The cations persist for quite some time after the light source is removed. Acids of the form $H^+SbF_5X^-$, where X is a halogen, also have been used to catalyze epoxy reactions.[22]

Summary

Epoxy structural adhesives have proved to be versatile and reliable compositions. Their widespread use and acceptance is in part due to the varied chemistry of the epoxy ring and the skill of scientists and adhesive formulators in developing high-quality compositions that can produce reliable, reproducible structural joints even when applied by relatively unskilled users. New compositions and chemical reactions continue to be disclosed, and it is certain that the knowledge of epoxy chemistry will continue to grow. Many of the new discoveries will find their way into new epoxy structural adhesives.

URETHANE STRUCTURAL ADHESIVES

Introduction

The term "urethane adhesive" as it is generally used encompasses a lot of chemistry that is not necessarily urethane chemistry. "Urethane" is the common name for the compound ethyl carbamate. In common usage, "urethane adhesive" generally means an adhesive polymer derived from isocyanate chemistry and reactions of isocyanates with active hydrogen compounds. However, isocyanate reactions do not always lead to urethane linkages, and there are ways of arriving at urethane linkages without involving isocyanates. In this section the common approach will be taken; that is, an adhesive that uses reactions of the isocyanate group to bring about polymerization in the bond line will be considered a urethane adhesive.

The study of industrial applications of isocyanate chemistry and polymers derived from isocyanates received much attention in Europe, particularly Germany, in the 1930s and during World War II. Patents on aspects of urethane chemistry appeared as early as 1937.[23] The effort going into understanding isocyanate chemistry and commercializing urethane products continues.[24,25]

The isocyanate group consists of a linear arrangement of nitrogen, carbon, and oxygen atoms, —N=C=O. Several possible electronic configurations can be drawn, most of which involve a positive charge on the central carbon atom. This partial positive charge on the carbon atom accounts for much of the reactivity of the isocyanate group with nucleophilic groups. Not only does the isocyanate group react with a variety of potential curing agents, but also it is very reactive with many of the adherend surfaces on which urethane adhesives are used.

The same high reactivity that makes the isocyanate a desirable reactant for structural adhesives also renders the unreacted isocyanate more acutely toxic than, for instance, epoxy adhesives. Only a few isocyanates are safe enough and easy enough to handle that they find widespread use in urethane structural adhesives.

Isocyanate Preparation

Organic isocyanates are the major building blocks of urethane structural adhesives. They can be synthesized by a variety of routes, but most of the commercially available isocyanate compounds used in adhesives are made by the reaction of a precursor primary amine or amine salt with phosgene, followed by dehydrohalogenation. The reaction with phosgene usually is carried out at a relatively low temperature, less than 60°C, and then the temperature is raised to 100 to 200°C to remove the HCl.

$$\underset{\text{amine}}{RNH_2} + \underset{\text{phosgene}}{COCl_2} \longrightarrow$$

$$\underset{\text{intermediate}}{[RNHCOCl + HCl]} \xrightarrow{\text{heat}} \underset{\text{isocyanate}}{RNCO} + 2HCl$$

Isocyanate synthesis by phosgenation

There are many possible side reactions, and the yield depends on the reaction conditions, which usually are specific to the starting materials and desired product. Much work has gone into determining the reaction conditions for manufacturing as clean a product as possible.

The amount of isocyanate in a commercial isocyanate or isocyanate containing formulation usually will be specified as weight

percent isocyanate (as NCO) or as an amine equivalent weight, which is the weight of material containing sufficient isocyanate to react with one mole of amine hydrogen.

Isocyanate Reactions

Organic isocyanates can undergo a large number of reactions with active hydrogen compounds. One test commonly used to determine the presence of active hydrogen atoms and the number of active hydrogens per molecule is the Tschugaeff-Zerewitinoff analysis or, more commonly, the Zerewitinoff test. An excess of a Grignard reagent, methyl magnesium iodide, is added to the sample to be tested, and the amount of methane evolved in measured.

$$CH_3MgI + RH \rightarrow CH_4 + RMgI$$

Zerewitinoff reaction

Urethane adhesives take their name from the product of the most common step-growth polymerization reaction used to generate the adhesive polymers. Isocyanates react with hydroxyl groups to create urethane (or carbamate) linkages:

$$RNCO + R'OH \longrightarrow R{-}\underset{H}{N}{-}\overset{\overset{\displaystyle O}{\|}}{C}{-}OR'$$

Urethane formation

If both the isocyanate and the hydroxyl-containing material are difunctional, if the mixture is made up to have one isocyanate per hydroxyl, if there are no side reactions, and if the reaction can be driven to completion, a single linear thermoplastic polymer should result. If one or more of the reactants is more than difunctional, it is possible to create an infinite three-dimensional network.

In general, primary hydroxyl groups are faster to react than secondary hydroxyls, which are in turn faster than tertiary hydroxyls, absent catalysts, and stearic factors. When reacted with hydroxyl-containing compounds, aliphatic isocyanates tend to be more sluggish than their aromatic counterparts. Urethane linkages made with tertiary hydroxyls tend to be less stable and at high temperature can dissociate into an olefin and an amine with loss of carbon dioxide. The isocyanate–hydroxyl reaction is catalyzed by metal compounds, particularly tin compounds such as dibutyl tin dilaurate, and by various amines.

Isocyanates will react with amines to produce substituted ureas, primary amines being faster than secondary amines.

$$R'NH_2 + RNCO \longrightarrow R'HN{-}\overset{\overset{\displaystyle O}{\|}}{C}{-}NHR$$

Urea formation from 1° amine

$$R'R''NH + RNCO \longrightarrow R'R''N{-}\overset{\overset{\displaystyle O}{\|}}{C}{-}NHR$$

Urea formation from 2° amine

The ureas are more rigid linkages than the urethane structure, but they also are generally more resistant to heat and chemical degradation. The reaction of isocyanates with amines is generally so rapid that it is nearly impossible to control the reaction well enough to make it useful in formulating urethane adhesives. What usually happens is that the reaction takes place faster than it is possible to mix the adhesive. If it is necessary to include urea linkages in the final production, that can be done by making a prepolymer which has the urea linkages in it already or by taking advantage of the slower reaction of isocyanates with water.

Small (much less than stoichiometric) amounts of diamines sometimes are added to the hydroxyl portion of the adhesive to provide a rapid but limited molecular weight increase as soon as the components are mixed.[26,27] In this way a two-package adhesive can be made that will flow easily before mixing but will not readily flow after mixing. Such adhesives can be applied to vertical surfaces

or overhead, and will remain in place until the bonds are closed and cured.

The reactions of amines and isocyanates are important in adhesives because of the possible reactions of isocyanates with water. Because isocyanates react readily with water, raw materials used in formulating adhesives must be dry, and the compositions must be protected from moisture, including atmospheric humidity, during storage. The first reaction with water is the formation of a carbamic acid, which rapidly loses carbon dioxide to form a primary amine.

Water NCO reaction

The primary amine then can react with another isocyanate to produce a urea. Thus one mole of water consumes at least two moles of isocyanate, builds molecular weight, and liberates carbon dioxide in the process. If this happens in a closed container such as a drum of adhesive, the result can be explosive, particularly as the reaction mass rises in temperature because of the exothermic reactions. On the other hand, these reactions can be useful in a bond line because under the proper conditions desirable urea linkages can be introduced into the curing adhesive through the reactions of small amounts of water normally present on the adherend surfaces.

The hydrogen atom attached to the nitrogen atom of the urethane group is active enough that it can react with another isocyanate group to produce an allophanate. This is an additional cross-linking mechanism for urethane polymers and can disturb the stoichiometry of the system by consuming an additional isocyanate group for each allophanate formed. Elevated temperatures usually are needed to produce allophanates in uncatalyzed systems; allophanation reactions can be catalyzed by tertiary amines.

Allophanate formation

In a similar fashion, a urea hydrogen atom can react with an additional isocyanate group to produce a biuret.

Biuret formation

Isocyanate groups also will react with themselves to form a variety of compounds. Two isocyanate groups can react to form a dimer or uretidinedione.

Isocyanate dimerization

These dimers can be dissociated to regenerate the original isocyanates with heat, and some of the dimers, such as the dimer of toluene diisocyanate, have become commercially important.

Three isocyanate groups can react to form a trimer or substituted isocyanurate ring. This reaction can be catalyzed by phosphines or strong bases such as socium acetate or sodium formate. The isocyanurate ring is thermally stable, had good chemical resistance, and can enhance the resistance of a urethane adhesive to aggressive environments.

Isocyanurate formation

Isocyanates also can react with each other to produce carbodiimides with the loss of carbon dioxide. This reaction requires high temperatures unless catalyzed by specific phosphorus compounds. Formation of carbodiimides normally is not an important cross-linking mechanism in polyurethane adhesives. However, carbodiimides are sold by Union Carbide (Ucarlnk™) and have been recommended as water scavengers and stabilizers for polyurethanes. The carbodiimide can react with water to give a urea, which still can react with additional isocyanate to produce a biuret.

$$R{-}NCO + R{-}NCO \xrightarrow[\text{or heat}]{\text{catalyst}} R{-}N{=}C{=}N{-}R + CO_2$$

Carbodiimide formation

$$R{-}N{=}C{=}N{-}R + H_2O \longrightarrow R{-}\underset{H}{N}{-}\overset{O}{\overset{\|}{C}}{-}\underset{H}{N}{-}R$$

Carbodiimide reaction with water

Carbodiimides can react with additional isocyanate groups to form uretone imines, which sometimes are used to modify polyisocyanates used in urethane structural adhesives.

$$R{-}N{=}C{=}N{-}R + R{-}NCO \longrightarrow$$

Uretone imine formation

Important Isocyanates

Dozens of isocyanate functional compounds have been synthesized, but only a few find much use in urethane structural adhesives. The choices are largely dictated by a combination of performance, price, and safety considerations. Most of the materials used in adhesives are derived from the aromatic isocyanates, toluene diisocyanate (TDI) and diphenylmethane diisocyanate (MDI).

2,4-Toluene diisocyanate

2,6-Toluene diisocyanate

(TDI)

Diphenylmethane diisocyanate
4,4-Methylene bis(phenyl isocyanate) (MDI)

Where color stability of the adhesive is important, and cure speed or cost is less important, aliphatic isocyanates are derived from isophorone diisocyanate (IPDI), hexamethylene diisocyanate (HDI), or 4,4'-dicyclohexylmethane diisocyanate ($H_{12}MDI$).

Isophorone diisocyanate (IPDI)

$$OCN{-}CH_2(CH_2)_4CH_2{-}NCO$$

Hexamethylene diisocyanate (HDI)

4,4'-Dicyclohexylmethane diisocyanate ($H_{12}MDI$)

The common commercial TDI is an 80/20 mixture of the 2,4- and 2,6- isomers. The pure

2,6- isomer is available also and is sometimes called TDS. To make toluene diisocyanate, toluene first is nitrated to produce a mixture of the 2,4- and 2,6- dinitro isomers. The dinitrate is reduced to the diamine and reacted with phosgene, which is followed by dehydrohalogenation to give the diisocyanate. Because of its relatively high vapor pressure and toxicity, adhesives rarely contain toluene diisocyanate monomer. Typically, excess TDI is reacted with another material such as trimethylol propane or a polyester diol or polyether diol to produce a higher-molecular-weight isocyanate functional compound, which is safer and easier to handle than the free TDI. Still, there will be some free TDI present in the adduct, and much work has gone into finding synthesis schemes to minimize the free TDI in prepolymers and adducts.[28]

Diphenyl methane diisocyanate or 4,4′-methylene bis (phenyl isocyanate), commonly called MDI, can be considered the first member of a series of polyisocyanates of the general form:

Aromatic polyisocyanates

and consequently the diisocyanate are produced in greatest yield, with nearly all of this being the 4,4′ isomer.

MDI is a solid at room temperature with a melting point of about 38°C, which usually is stored and shipped in the molten state for convenience. However, on standing in the liquid state, the MDI slowly dimerizes, and the liquid MDI becomes saturated with dimer at about 1 percent dimer by weight. The dimer then begins to precipitate, and the isocyanate content of the liquid falls because isocyanate is consumed in the dimer formation and removed in the precipitation.

In general, the aliphatic isocyanates have better resistance to degradation by ultraviolet light than do the aromatic isocyanates. In spite of their higher cost and lower reactivity, the aliphatic isocyanates do find some use in adhesives where the adhesive is visible and color stability is important in the final assembly.

Hexamethylene diisocyanate (HDI) can be used as an aliphatic cross-linker in urethane adhesives but presents significant health risks because of its toxicity and high vapor pressure. HDI is more commonly used in the form of its biuret, which is much safer and easier to handle than HDI.

Biuret of hexamethylene diisocyanate

MDI is the most important member of the series although materials with n of 1 or more also are available. The precurser amine is made by condensation of aniline hydrochloride with formaldehyde, followed by reaction with phosgene and dehydrohalogenation. When the aniline is present in excess, the diamine

Isophorone diisocyanate (IPDI) is another low-viscosity aliphatic diisocyanate that is useful in formulating light-stable polyurethane adhesives, and is somewhat lower in cost than hexamethylene diisocyanate. IPDI has a low vapor pressure at room temperature but, like most low-weight isocyanates, still can present

a health hazard. Higher-molecular-weight adducts of IPDI, such as its trimer, the isocyanurate of IPDI, are available. The trimer is a high-melting solid (100–115°C), and one loses the advantage of the low viscosity liquid in using it but gains in safety and easy handling.

IPDI is a cycloaliphatic diisocyanate, but the two isocyanate groups are not equivalent. One is attached directly to the aliphatic ring, and the other is separated from the ring by a methylene group. Because their environments are different, the reactivities of the two groups are different; and the reactivity can depend on the choice of catalyst. For instance, Ono and coworkers[29] have shown that when IPDI is reacted with a primary hydroxyl group, the primary NCO is most reactive when the catalyst is 1,4-diazabicyclo [2.2.2] octane. The secondary NCO attached to the ring is the most reactive when dibutyl tin dilaurate is the catalyst.

The saturated analog of MDI, 4,4′ dicyclohexyl methane diisocyanate, has found limited used as an aiphatic isocyanate in adhesives. This material is known by a variety of names including Desmodur W™ (Mobay), hydrogenated MDI (or HMDI or H_{12} MDI), reduced MDI (RMDI), and saturated MDI (SMDI). It is a low viscosity liquid with a fairly high vapor pressure; so it too must be handled with care. In adhesive compositions, the diisocyanate usually is used to make an isocyanate functional prepolymer by reacting excess diisocyanate with a hydroxyl or amine functional polymer such as a polyester diol.

Because it contains two saturated six-member rings, 4,4′ dicyclohexyl methane diisocyanate can exist in three isomeric forms, in which the orientations of the NCO groups with respect to the rings and each other are different. The *trans–trans* isomer is a solid at room temperature so that over time this isomer can precipitate from the remaining liquid isomers. Samples of 4,4′ dicyclohexyl methane diisocyanate that have been stored at cool temperatures for a while often have a solid layer of *trans–trans* precipitate in the bottom.

Two examples of a new family of diisocyanates have been offered by American Cyanamid, the *meta* and *para* isomers of tetramethyl xylene diisocyanate (TMXDI).[30]

The *para* isomer of tetramethyl xylene diisocyanate is a solid melting at 72°C and so has received less attention in adhesives than has the liquid *meta* isomer (m.p. −10°C). These isocyanates are interesting because they contain an aromatic ring, but the NCO groups themselves are aliphatic isocyanates and have reaction characteristics typical of aliphatic diisocyanates. They are even somewhat more sluggish in reacting than the more standard aliphatic isocyanates, making the reactions easier to control. Compounds such as dimethyl tin dilaurate, lead octoate, or tetrabutyl diacetyl distannoxane have been shown to be effective catalysts for the isocyanate–hydroxyl reaction. The manufacturers claim that although they are still toxic, the *p*-TMXDI and *m*-TMXDI appear to be less toxic than many other isocyanates.

m-Tetramethyl xylene diisocyanate (*m*-TMXDI)

Blocked Isocyanates

Blocked isocyanates are compounds formed by the reaction of an organic isocyanate with an active hydrogen compound where the reaction is reversible with heat.The blocked isocyanate can be used in formulating adhesives or other reactive compositions even in the presence of materials that normally would react rapidly with the isocyanate. The isocyanate can be regenerated by heating the mass to a temperature at which the blocked isocyanate begins to dissociate at an acceptable rate. Phenol is one example of a blocking agent used with isocyanates. There are many other blocking agents in use, including lactams, oximes, and malonates. Even isocyanate dimers such as TDI dimer could be considered blocked isocyanates because they will dissociate

with heat to regenerate the isocyanates.

Phenol blocking reaction

There are two major limiting factors to the use of blocked isocyanates in urethane structural adhesives. First, the adherends must be sufficiently heat-resistant to withstand the temperatures needed to cause rapid dissociation of the blocked isocyanate. Second, and perhaps more important, there usually must be some mechanism for removing the blocking agent from the bond line. Otherwise the blocking agent is present to continue competition with the intended curing agent for reaction with the isocyanate. After the adhesive is cured, blocking agent remains trapped in the bond line and could contribute to poor mechanical properties in the bond or poor resistance to harsh environments. The need to remove the blocking agent makes the blocked isocyanates particularly attractive in applications such as laminating where a large exposed area of adhesive is cured before the bond is closed.

Summary

Urethane structural adhesives have proved quite successful in bonding applications that take advantage of their toughness, particularly when the adhesive must function at low temperatures. The reactivity of the isocyanate group may require some extra care in handling and storing the adhesives but the same reactivity provides many opportunities for the polymer chemist and adhesive formulator. The principal reaction used in curing urethane structural adhesives is the reaction of the isocyanate group with primary and secondary hydroxyl groups. Both the isocyanate groups and the hydroxyl groups can be carried on a wide variety of low molecular weight oligomers such as polyesters, polyethers, polybutadienes, polyurethanes, and polymer blends so that many of the desired properties of the final adhesive can be built into the polymer before the adhesive is cured.

ACRYLIC STRUCTURAL ADHESIVES

Introduction

Acrylic structural adhesives are relatively new products compared to epoxy or urethane adhesives. They first appeared in Europe in the mid-1960s and were commercialized in the United States a few years later. The first offerings were rather brittle products of use in limited applications where toughness and flexibility were not critical. However, because these new acrylic structural adhesives polymerized in the bond line through a free radical chain growth polymerization mechanism, they did have the big advantage of curing rapidly at room temperature. This polymerization mechanism offered a variety of potential advantages in manufacturing. Bonded structures could be assembled relatively rapidly without the need of curing ovens. Assemblies did not need to be stored for long periods while room-temperature cures were completed. The adhesives could be made at very low viscosities for easy handling and dispensing. In the past two decades there has been much creative work done with the chemistry of acrylic adhesives so that now they are available in many forms with a wide range of properties. Figure 20.2 shows the spray application of an acrylic structural adhesive to fiber-reinforced plastic panels that are assembled into farm silos in Japan.

Acrylic Monomers

All acrylic structural adhesives consist basically of a solution or a mixture of polymers and

Fig. 20.2. Spray application of an acrylic adhesive to fiber-reinforced plastic panels, which are assembled into farm silos in Japan. (*Courtesy Lord Corp., Erie, PA.*)

unsaturated, low molecular weight, free-radical-polymerizable monomers with other materials added as needed for the particular intended use. In addition, some precursor of the polymerization initiator will be present. The most common monomers used are methacrylic acid (2-methyl propenoic acid) and its esters or, less commonly, acrylic acid (propenoic acid) and its esters.

Methacrylic acid

Acrylic acid

Several processes for making acrylic acid have been developed.[31] Most manufacturers now use one of two processes, the oxidation of propylene or the Reppe process, which involves the oxidation of acetylene.

The oxidation of propylene is a two-step process with specific catalysts used in each step. In the first, the propylene is converted to acrolein, which then is further oxidized to acrylic acid in the second step. The acrylic acid then can be esterified with the appropriate alcohols.

In the Reppe pocess, acetylene is reacted with carbon monoxide, nickel carbonyl, and water to produce acrylic acid. If the appropriate alcohol is used in place of the water, the acrylate ester can be produced directly. Acrylates made by the Reppe process tend to contain more propionates than do those made

by oxidation of propylene.

Reppe process for making acrylic acid

If carbon monoxide is supplied as a separate feed in the presence of small amounts of nickel carbonyl, the reaction also will proceed with the nickel functioning as a catalyst.

Methacrylic acid can also be produced by a variety of processes including oxidation of ethylene, propylene, and isobutylene.[32] The most common commercial process for making methacrylic acid is the acetone cyanohydrin (ACN) process. The feedstocks for this process are acetone, hydrogen cyanide, and sulfuric acid. The acetone and HCN are reacted under alkaline conditions to produce the cyanohydrin. Reaction of the cyanohydrin with sulfuric acid ultimately produces methacrylamide sulfate, hydrolysis of which produces methacrylic acid. Methyl methacrylate can be made either by esterifying the acid or directly by reacting the amide sulfate with methanol.

Cyanohydrin

Methyl methacrylate

Methacrylic acid

The choice of monomers that are useful in acrylic structural adhesives is rather limited. Cost always is an important factor, and because acrylic structural adhesives consume only a very small portion of the world's output of acyrlic monomer, the formulator usually must rely on acrylic monomers which are made in large quantity for other uses. In addition, the monomers must polymerize readily at room temperature. If a mixture of monomers is to be used, the monomers must copolymerize easily. Finally, the monomers must be good solvents or dispersants for the polymers used in formulating the adhesive. The acrylic monomers finding most use in acrylic structural adhesives are methacrylic acid and its esters, particularly methyl

methacrylate. These two usually make up the bulk of the monomer present.

Curing

The curing reaction of an acrylic structural adhesive is the chain-growth polymerization of the acrylic monomer. The monomer units are not reactive with each other but react only with a growing chain having an active site on one end. In order to begin a chain, one must generate an initiator in the monomer solution. In the case of acrylic structural adhesives, this initiator nearly always is a free radical, a species having an unpaired, reactive electron.

In addition, a growing chain might be terminated by reaction with a variety of unknown impurities invariably present in any mixture.

The initiator usually is created in a redox reaction that follows a free radical pathway. Common reactions involve the reduction of an organic peroxide by some reducing agent such as an amine or an ion capable of undergoing a one-electron transfer reaction. One example of an efficient free radical initiator generating reaction is the reduction of diacyl peroxides such as benzoyl peroxide (BPO) by tertiary aromatic amines such as N,N-dimethyl aniline (DMA).

Benzoyl peroxide reaction with N,N-dimethyl aniline

After chains have been initiated, there are three general types of reaction that can occur: propagation, chain transfer, and chain termination. If "*" represents the active site, "A*" represents the initiator, and M is a monomer unit, then the four reaction types are:

Initiation: $A^* + M \rightarrow AM^*$
Propagation: $AM^* + nM \rightarrow AM_nM^*$
Chain transfer: $AMnM^* + M \rightarrow AM_nM + M^*$
Termination: $RM^* + R'M^* \rightarrow RMMR'$ (combination)
$RM^* + R'M^* \rightarrow RMC{=}C + RMCC$ (disproportionation)

Peroxyesters can be reduced by metal ions such as copper or iron to produce free radicals that will initiate chain growth polymerization.

Reduction of a peroxyester

Free radical initiators also can be generated by thermal decomposition of certain compounds such as organic peroxides or azo compounds. Ultraviolet light or electron beams can initiate the polymerization of acrylic monomers.

The chain-growth polymerization, once it has been initiated, proceeds in two fairly distinct stages. The important point is that the monomers are not reactive with each other but only with the growing chains. Over its lifetime one propagating free radical could produce a chain many monomer units long. This chain is one single polymeric molecule dissolved in the remaining monomeric liquid. As free radicals are produced and propagate, more polymer "solute" chains are produced at the expense of the monomer "solvent." Eventually there will remain too little monomer to dissolve all the polymer being formed.

The point at which there is too little monomer to be an effective solvent for the acrylic polymer forming in the polymerization reaction (and also any oligomeric materials added to the formulation) can occur very early in the polymerization reaction. It is possible for the polymer to precipitate as a totally separate single phase, leaving the liquid monomer behind, but this is undesirable. Usually enough polymer forms that the polymer chains become entangled, and a gel saturated with monomer is formed. This is significant from two aspects. Once the gel has formed, the reacting mass in the bond line should not be disturbed until the cure is much further advanced because the composition may no longer be fluid enough to re-form the bond. Second, the rules of the polymerization process change with the onset of the Trommsdorf effect (or "gel" effect), and the remaining polymerization becomes much more rapid.

Prior to the gel point, the growing polymer chains with the reactive terminal sites are mobile and have easy access to monomer and to each other. Both chain growth and termination are significant reactions. After the gel point, the bimolecular termination reactions become much less significant because the chains no longer are freely mobile, and the probability of two radicals coming close enough to interact greatly decreases. The low molecular weight monomer still is able to diffuse to the reactive sites, so chain growth continues. The average life of the active radical site on the growing chain becomes much longer after the gel than it was before so that the overall polymerization proceeds at a much faster rate after the gel.

After the free radical initiators are generated in the adhesive, by mixing the redox couple, ultraviolet light, heat, or other means, the adhesive remains liquid for some period of time until the gel point is reached. Then the molecular weight, strength, and toughness of the adhesive build rapidly.

One difficulty sometimes encountered with free radical polymerizations in acrylic adhesives is the inhibition of the reaction by oxygen. Oxygen contains two unpaired electrons and readily attaches to the propagating free radical on a growing chain. The resulting oxy radical is much less reactive than the original carbon radical, and the chain growth effectively stops. Even if it proceeds, the chain backbone will contain a less stable peroxide linkage. The result is that the surface of acrylic adhesives cured in the presence of oxygen often will be low molecular weight and will be tacky even though the bulk of the adhesive is fully cured.

Formulation

Although the free radical chain growth polymerization brings processing advantages, it also brings limitations. The product of the polymerization of the monofunctional acrylic monomers is a linear, thermoplastic polymer. In the case of a copolymer that is largely methyl methacrylate, the glass trasition temperature of the polymer will be well above room temperature if the polymer has been driven to a high-enough molecular weight. The resulting adhesive is brittle, and although its shear strength might be quite high, its ability to withstand peeling forces or impacts will be low. One could use monomer mixtures yielding copolymers having lower glass transition temperatures and expect that the adhesive might be somewhat tougher. However,

this parameter is difficult to control because in curing the acrylic adhesive in the bond line, one really is doing a bulk polymerization under uncontrolled conditions and trying to produce a specific polymer with a specific set of material and performance properties. The approach taken by most adhesive manufacturers or formulators is to use "prepolymers," high molecular weight oligomers that can be made under controlled conditions and then dissolved or dispersed in the acrylic monomer. By forming key elements of the polymer chains under controlled conditions, many of the desired properties of the final product can be built into the adhesive when it is formulated and before the material is cured.

The simplest approach to improving toughness is to dissolve or disperse a nonreactive, rubbery polymer in the monomer mixture that is to be polymerized. Examples of such materials are nitrile rubbers, polyethers, and acrylic rubbers. The choice will depend on cost, desired properties, and the solubility of the polymer in the monomer. As the monomer polymerizes, it will lose its solvating ability for the dissolved polymer. Eventually, when enough of the monomer has been consumed, it no longer can hold the polymer in solution, and the polymer precipitates. The polymer then can segregate as a totally separate phase or as distinct domains heterogeneously dispersed throughout the acrylic polymer matrix, or it can be homogeneously trapped in the acrylic matrix, depending on the rates of reaction, the compatibility with the acrylic polymer, and the relative mobility of the dissolved polymer.

One way of influencing the way in which the added oligomer is distributed in the final cured produce is to provide reactive sites on the oligomer so that it can be incorporated into the acrylic matrix as it forms. For instance, if the added oligomer contains terminal, active, acrylic, or methacrylic unsaturation, it can be easily incorporated into the growing acrylic polymer chains as the adhesive cures. Then it is unable to precipitate as a separate phase and must remain more or less uniformly distributed throughout the matrix. On the other hand, incompatible polymers can be added to deliberately encourage the formation of reinforcing domains. If the oligomer has two or more functional groups capable of participating in the polymerization, there is at least the possibility of forming cross-links in the polyacrylate matrix.

A technology developed at Du Pont[32,33] combines the use of reactive sites on the oligomers with the initiation reaction. The resulting family of acrylic structural adhesives has become popularly known as "second generation acrylics." They consist essentially of solutions of chlorosulfonated polyethylene (Du Pont Hypalon™) in acrylic or methacrylic monomers. The sulfonic acid groups present on the polymer will react with amine aldehyde condensation products to generate free radical initiators, at least some of which are claimed to be on the oligomer backbone. The claim is that the adhesive polymer chains then grow by graft polymerization from various sites on the rubber oligomer. Peroxides and metal ions can be added to accelerate the initiation and polymerization. One of the significant advantages of this type of acrylic structural adhesive is the ability to bond oily or dirty metals with only minimal surface preparation, although some surface preparation usually is necessary to improve bond durability.

Other acrylic structural adhesives can be used on poorly prepared metal surfaces. Inclusion of reactive monomers having phosphate groups can lead to significant improvements in primary bonds to many unprepared metals.[34]

Another innovative approach to initiating the polymerization of acrylic structural adhesives is illustrated by a family of adhesives that uses the adherend surface itself as part of the chemical reaction that generates the free radical initiator.[35,36] The compositions contain an acidic material, a sulfonyl halide group, and a transition metal compound. Upon contact with a metal surface such as iron, zinc, copper, cadmium, and their alloys, the components react with the metal surface through a free radical pathway, producing effective initiators for the polymerization of the acrylic monomers. Later extensions of this concept provide two-part adhesives employing

the same initiation chemistry,[37] in which a metal powder is included in one portion of the adhesive.

Summary

Acrylic structural adhesives have the advantage of easy handling and processing and rapid cure at room temperature. The ability of some compositions to adhere to unprepared metal surfaces can allow substantial cost savings in manufacturing processes. The brittleness of the polymers made from the monomers often can be overcome by formulating to include tough, reactive oligomers in the liquid adhesive. The variety of possible initiators and mechanisms for generating them has brought many innovative compositions to market.

HYBRID ADHESIVES

Adhesives usually are classified as epoxy, urethane, or acrylic, based upon the chemistry used to bring about the olymerization of the liquid adhesive. However, the distinction among the various types of adhesives is not always clear, and many hybrid adhesives have been developed and marketed. The objective usually is to take advantage of the desirable mechanical properties or chemical resistance of one polymer while retaining the processing attributes of a different cure system.[38]

It is common in acrylic structural adhesives to use oligomers that have a desirable backbone and are terminated with free-radical-polymerizable bonds. A variety of isocyanate-terminated polyurethanes can be adapted to use in acrylic structural adhesives by reacting the terminal isocyanates with a hydroxy functional acrylic monomer such as 2-hydroxyethyl methacrylate.[34,39,40]

Sometimes the system is formulated so that multiple cure mechanisms are possible and can occur sequentially or simultaneously. Compositions that rely on both epoxy and urethane chemistry are examples.[27,41,42] These are compositions containing the diglycidyl ether of bisphenol A, an isocyanate or isochanate-terminated prepolymer, amines or other reactants for either epoxy or isocyanates, and catalysts.

CONCLUSION

Epoxy, urethane, and acrylic structural adhesives have been commercially successful because each can be used, under the proper conditions, to make reliable, durable, and useful adhesively joined assemblies. The adhesives are classified according to the polymerization reactions used to bring the liquid adhesive to a high-molecular-weight load-bearing state. The reactive sites on the uncured adhesive and the overall polarity of the cured adhesive are important in the adhesion of the polymer to any specific material. The polymerization reactions determine the processing requirements and are important to the mechanical properties and environmental resistance of the cured adhesive. Through the use of specific reactive oligomers it is possible to build certain molecular structures into the composition that will remain in the cured adhesive to provide desired mechanical or chemical properties.

The successful use of adhesive joining in producing any assembly depends on viewing the assembly as a whole from the very beginning of its design. The key is remembering that the adhesive is only one component of the assembly. Adhesion is a property of the whole assembly.

REFERENCES

1. Patrick, R. L. (Ed.), *Treatise on Adhesion and Adhesives, Vol. 1, Theory*, Marcel Dekker, New York, 1967.
2. Patrick R. L. (Ed.), *Treatise on Adhesion and Adhesives, Vol. 2, Materials*, Marcel Dekker, New York, 1969.
3. Kaelble, D. H., *Physical Chemistry of Adhesion*, Wiley-Interscience, New York, 1971.

4. Patrick, R. L. (Ed.), *Treatise on Adhesion and Adhesives, Vol. 3*, Marcel Dekker, New York, 1973.
5. Skeist, I. (Ed.), *Handbook of Adhesives*, 2nd ed., Van Nostrand Reinhold, New York, 1977.
6. Wake, W. C., *Adhesion and the Formulation of Adhesives*, 2nd ed., Applied Science, New York, 1982.
7. Lee, L-H., *Adhesive Chemistry, Developments and Trends*, Plenum Press, New York, 1984.
8. Panek, J.R., and Cook, J.P., *Construction Sealants and Adhesives*, 2nd ed., Wiley-Interscience, New York, 1984.
9. Adams, R. D., and Wake, W. C. *Structural Adhesive Joints in Engineering*, Elsevier, New York.
10. Hartshorn, S. R. (Ed.), *Structural Adhesives Chemistry and Technology*, Plenum Press, New York, 1986.
11. Kinlock, A. J., *Adhesion and Adhesives Science and Technology*, Chapman and Hall, London, 1987.
12. Mark, H. F., "Future Improvements in the Cohesive and Adhesive Strength of Polymers—Part I," *Adhesives Age*, **22** (7), 35–40 (1979).
13. Mark, H. F., "Future Improvements in the Cohesive and Adhesive Strength of Polymers—Part II," *Adhesives Age*, **22** (9) 45–50 (1979).
14. Lee, H., and Neville, K., *Handbook of Epoxy Resins*, McGraw-Hill, New York, 1967.
15. Castan, P., "Process of Preparing Synthetic Resins," U.S. Patent No. 2,324,483, 1943.
16. Graham, J, A., and O'Connor, J. E., "Epoxy with Low-Temperature Cure and High-Temperature Properties Developed," *Adhesives Age*, **21** (7), 20–23 (1978).
17. Crivello, J. V., "Heat Curable Compositions," U.S. Patent No. 4,173,551, 1979.
18. Crivello, J. V., "Heat Curable Cationically Polymerizable Compositions and Method of Curing Same with Onium Salts and Reducing Agents," U.S. Patent No. 4,216,288, 1980.
19. Crivello, J. V., "UV Curable Compositions and Substrates Treated Therewith," U.S. Patent No. 4,319,974, 1982.
20. Crivello, J. V., and Ashby, B. A., "Methods of Adhesive Bonding Using Visible Light Cured Epoxies," U.S. Patent No. 4,356,050, 1982.
21. Crivello, J. V., and Lee, J. L., "Photocurable Compositions," U.S. Patent No. 4,442,197, 1984.
22. Tarbutton, K. S., and Robins, J., "Acid Catalyzed, Toughened Epoxy Adhesives," U.S. Patent No. 4,846,905, 1989.
23. Bayer, O., German Patent No. 728.981, 1937.
24. Saunders, J. H., and Frisch, K. C., *Polyurethanes, Chemistry and Technology, Vols. 1 and 2*, Interscience, New York, 1962, 1964.
25. Oertel, G. (Ed.), *Polyurethane Handbook*, Hanser, Munich, 1985.
26. Fabris, H. J., Maxey, E. M., and Uelzmann, H., "Urethane Adhesive Having Improved Sag Resistance," U.S. Patent No. 3,714,127, 1973.
27. Goel, A. B., "Sag Resistant Urethane Adhesives with Improved Antifoaming Property," U.S. Patent No. 4,728,710, 1988.
28. Baueriedel, H., "Adhesives Based on Polyurethane Prepolymers Having a Low Residual Monomer Content," U.S. Patent No. 4,623,709, 1986.
29. Ono, H-K., Jones, F. N., and Pappas, S. P., "Relative Reactivity of Isocyanate Groups of Isophorone Diisocyanate. Unexpected High Reactivity of the Secondary Isocyanate Group," *Journal of Polymer Science: Polymer Letter Edition*, **23**, 509–515 (1985).
30. Arendt, V. D., Logan, R. E., and Saxon, R., "*m*- and *p*-TMXDI: Two New Isocyanates for the Polyurethane Industry," *Journal of Cellular Plastics*, pp. 376–383 (Nov./Dec. 1982).
31. Nemec, J. W., and Bauer, W. Jr., "Acrylic Acid and Derivatives," in *Encyclopedia of Chemical Technology, Vol. 1*, Martin Grayson (Ed.), pp. 330–354, Wiley-Interscience, New York, 1978.
32. Nemec, J. W., and Kirch, L. S., "Methacrylic Acid and Derivatives," in *Encyclopedia of Chemical Technology, Vol. 15*, Martin Grayson (Ed.), pp. 346–376, Wiley-Interscience, New York, 1978.
33. Briggs, P. C., "Bonding Method Employing Novel Methacrylate Based Adhesive Compositions; Chlorinated or Chlorosulfonated Polyethylene as Adhesion Promoter," U.S. Patent No. 4,773,957, 1988.
34. Zalucha, D. J., Sexsmith, F. H., Hornaman, E. C., and Dawdy, T. H. "Structural Adhesive Formulations," U.S. Patent No. 4,223,115, 1981.
35. Daico, D. J., "Structural Adhesive Formulations," U.S. Patent No. 4,703,089, 1987.
36. Damico, D. J., and Bennett, R. M., "Structural Adhesive Formulations and Bonding Method Employing Same," U.S. Patent No. 4,855,001, 1989.
37. Damico, D. J., Mushrush, K. W., and Bennett, R. M., "Two-Part Adhesive for Bonding Plastics and Metals," U.S. Patent No. 4,857,131, 1989.
38. Pohl, E., and Osterholz, F. D., U.S. Patent No. 4,645,816, Feb. 24, 1987.
39. Brownstein, A. M., U.S. Patent No. 3,428,614, 1969.
40. Su, W-F., A., U.S. Patent No. 4,618,632, Oct. 21, 1986.
41. Hawkins, J. M., "Epoxy Resin Adhesive Compositions Containing an Isocyanate Terminated Polyurethane Prepolymer and a Chain Extender," U.S. Patent No. 3,636,133, 1972.
42. Trieves, R., and Pratley, K. G. M., U.S. Patent No. 4,623,702, Nov. 1986.

21

Manufactured Textile Fibers

Bhupender S. Gupta*†

TEXTILE BACKGROUND

The first conversion of naturally occurring fibers into threads strong enough to be looped into snares, knitted to form nets, or woven into fabrics is lost in prehistory. Unlike stone weapons, such threads, cords, and fabrics—being organic in nature—have in most part disappeared, although in some dry caves traces remain. There is ample evidence to indicate that spindles used to assist in the twisting of fibers together had been developed long before the dawn of recorded history. In that spinning process, fibers such as wool were drawn out of a loose mass, perhaps held in a distaff, and made parallel by human fingers. (A maidservant so spins in Giotto's *The Annunciation to Anne*, ca. A.D. 1306, Arena Chapel, Padua, Italy.[1] A rod (spindle), hooked to the lengthening thread, was rotated so that the fibers while so held were twisted together to form additional thread. The finished length then was wound by hand around the spindle, which, in becoming the core on which the finished product was accumulated, served the dual role of twisting and storing, and, in so doing, established a principle still in use today. (Even now, a "spindle" is 14,400 yards of coarse linen thread.) Thus, the formation of any threadlike structure became known as spinning, and it followed that a spider spins a web, a silkworm spins a cocoon, and manufactured fibers are spun by extrusion, although no rotation is involved.

It is not surprising that words from this ancient craft still carry specialized meanings within the textile industry and have entered everyday parlance, quite often with very different meanings. Explanations are in order for some of the words used in the following pages. For example, as already indicated, "spinning" describes either the twisting of a bundle of essentially parallel short pieces of wool, cotton, or precut manufactured fibers into thread or

*College of Textiles, North Carolina State University.
†The author dedicates this chapter to Dr. Robert W. Work, Professor Emeritus—a long-time friend and mentor—on the occasion of his eighty-fifth birthday. (*Editor's note*: Dr. Work authored this chapter in the sixth and seventh editions of this book and a co-author with Dr. Colin L. Brown in the 8th edition. The author gratefully acknowledges the assistance of associates both from within the College of Textiles and from outside, including several fiber producting companies. he is also indebted to Dr. Robert E. Work for critical review of the manuscript and many helpful suggestions.

the extrusion of continuous long lengths of manufactured fibers. In the former case, the short lengths are known as "staple" fibers, and the resulting product is a "spun yarn," whereas the long lengths are called "continuous filament yarn," or merely "filament yarn." Neither is called a "thread," for in the textile industry that term is reserved for sewing thread and rubber or metallic threads. Although to the layperson "yarn" connotes a material used in hand knitting, the term will be used in the textile sense hereinafter.

Before manufactured fibers are discussed, it is necessary to define some terms.* The "denier" of a fiber or a yarn defines its linear density, that is, the mass in grams of a 9000-meter length of the material at standard conditions of 70°F and 65 percent relative humidity. Although denier is actually a measure of linear density, in the textile industry the word connotes the size of the filament or yarn. Fibers usually range from 1 to 15 denier, yarns from 15 to 1650. Single fibers, usually 15 denier or larger, used singly, are termed "monofils." The cross-sectional area of fibers of identical deniers will be inversely related to their densities, which range from 0.92 g/cc for polypropylene to 2.54 g/cc for glass. Because by definition denier is measured at standard conditions, it describes the amount of "bone-dry" material plus the moisture regain, which ranges from zero for glass and polypropylene to 13 percent for rayon. It should be mentioned that some years ago scientific organizations throughout the world accepted the word "tex," this being the mass of one kilometer of the material, as a more useful term than denier. "Tex" is an accepted adjunct to the SI, or International System of Units, but it has received only limited acceptance in commerce, whereas the SI units are being employed increasingly in scientific organizations. Furthermore, the sizes of cotton, wool, and worsted yarns, and yarns containing manufactured fibers but produced by the traditional cotton, wool, or worsted systems, still are expressed in the inverse-count system that has been used for centuries.

*Each year the ASTM publishes in its *Book of Standards*, the most recent and accepted definitions and test methods used in the textile and fiber industries.

The "breaking tenacity" or more commonly, "tenacity," is the breaking strength of a fiber or a yarn expressed in force per unit denier, that is, in grams per denier, calculated from the denier of the original unstretched specimen.

Fig. 21.1. Force–elongation curves of natural and manufactured staple fibers at standard conditions of 70°F and 65 percent relative humidity.

Fig. 21.2. Force–elongation curves of manufactured textile continuous filament yarns at standard conditions of 70°F and 65 percent relative humidity.

"Breaking length" expresses the theoretical length of yarn that would break under its own weight, and is used mostly in Europe. "Elongation" means "breaking elongation" and is expressed in units of length calculated as a percentage of the original specimen length.

Typical force–elongation curves of some manufactured and natural staple fibers and textile-type manufactured filaments are shown in Figs. 21.1 and 21.2.

HISTORY

Early humans, over time, became aware of the presence and usefulness of fibrous materials available from such sources as seeds, leaves, stems, animal coverings, and cocoons. They learned to spin, weave, knit, felt, or braid these fibers to protect their bodies and improve their life-style. A few hundred years ago, it has been suggested, someone took a clue from a busy worm and thought that it should be possible for humans to make a silk-like fiber that would be of commercial value. Curiosity combined with simple experiments strengthened that premise and much later led to the development of viscose rayon, only partially a manufactured fiber, in the nineteenth century. Further curiosity, war-time need, and superior commitment by modern researchers led to the synthesis of a totally synthetic fiber in the 1930s. The knowledge gained and the spark thus ignited resulted in the development of other fibers to such an extent that manufactured fibers now dominate the market in the industrialized nations in all major categories of apparel, home furnishing, and industrial end uses.

The story of the development of manufactured fibers is of great historical interest, beginning in 1664 when Robert Hooke, an Englishman, suggested that it should be possible to make a fiber much like silk that could be of value in the market place. Andemars, a Swiss chemist, received the first patent for making silk-like fiber in 1855. He drew fibers by dipping a needle and pulling it out from a solution of cellulose nitrate containing some rubber.

The credit for using a spinnerette and forcing a solution through it for producing a fiber, however, goes to the English scientist Sir Joseph W. Swan, in the early 1880s. The first person to put the idea into commercial practice was the French chemist Count Hilaire de Chardonnet, who built the first plant to commercially produce a fiber based on regenerated cellulose, called "artificial silk," at Besancon, France, in 1891. During the last years of the nineteenth century and the beginning of the twentieth, progress was so rapid that the production of this fiber increased from several thousand pounds in 1891 to over two million in 1910. Commercial production of the fiber in the United States began in 1910 with the opening of the first plant. In 1924, the industry gave the "artificial silk" fiber a new name, rayon.

By that year, 1910, the brothers Camille and Henry Dreyfus had discovered a practical method for producing cellulose acetate polymer and were making plastic film and toilet articles in Basel, Switzerland. During World War I, they built a plant in England to produce acetate dope for painting airplane wings to render them air-impervious. The success of the product led the U.S. government to invite the Dreyfus brothers to build a plant in the United States, which started commercial production in 1924.

The successful manufacture of these two fibers, although based on fibrous materials available in nature, marked the beginning of the development of manufactured fibers in the 1930s. This effort, initiated by a technological breakthrough, was marked by the work of W. H. Carothers, aimed at learning how and why certain molecules joined to form large molecules, or polymers.[2] Fibers were described as being composed of high molecular weight linear polymers; and the first one to be manufactured, nylon 66, was synthesized and produced on a commercial scale in 1939. It was quickly followed by nylon 6, the second most widely used nylon, and modacrylic (1949), olefin (1949), acrylic (1950), polyester (1953), and triacetate (1954). Glass had joined this group of large-production items earlier, in 1936.

Several other fibrous materials have been produced, but they are regarded as fibers with

special performance characteristics, used either in limited textile or specialized industrial applications. Some of these worth noting are Spandex (1959), Aramid (1961), PBI (polybenzimidazole) (1983), and Sulfar (1983).

Thus, the period from the 1930s to the 1960s can be considered as a time of discoveries and innovations for manufactured fibers, when the majority of the basic fibers were developed. The years since then may be thought of as a period of modification of performance characteristics. The basic generic materials have been manipulated both chemically and physically to produce a wide variety of different fibers, tailored to secure the desired characteristics for specific-end-use products. Thus, fibers can be extruded in different shapes and sizes for special purposes. They can be modified to offer greater comfort, flame resistance, or static-free behavior in apparel; they can offer soil-release and other desirable characteristics for carpets; they can be developed with unique surface characteristics, easier dyeability, or better blending qualities. The industry has begun to discover many possibilities for modifying the behavior of a given fibrous materials. It has learned how to produce new fibers with greater strength, greater thermal resistance, or other special qualities.

Essentially, then, no new, large-volume, highly profitable fibers have been developed since the midfifties. Instead, the existing ones have become commodities with all the economic impact thereby implied. No major chemical engineering processes have been added, although the previously described ones have been modified to allow for spinning of liquid crystalline polymers or the formation of gel spun fibers. Research activity has been reduced and centered essentially on modifications of fiber size, shape and properties, and many variants now are successfully marketed. Production volumes have increased enormously for nylon, polyester, and polyolefin.

VOLUME OF PRODUCTION

Figure 21.3 compares population growth with the production of manufactured fibers and the mill consumption of natural fibers in the United States. Per capita consumption of all fibers, starting at a level in the 1920s of about 30 lb, rose to approximately 40 lb following World War II and reached a level at or about 50 lb in the 1970s. It topped 60 lb in 1973, dropped below 45 lb in 1982, and in 1991 was at 55 lb. But clearly overshadowing the increases resulting from population growth and a higher standard of living are the volumes produced of, first, the cellulosic manufactured fibers and, second, the noncellulosic or completely synthetic fibers. The consumption of manufactured fibers increased from about 2 billion lb in 1960 to nearly 9 billion lb in 1990. During the 1980s the increase was about 11 percent, made up of a decrease in cellulosics of about 200 million lb (25%) and an increase in synthetics of about 1.2 billion lb (17%). Of the 13.2 billion lb of fibers used in the United States in 1990, 4.6 percent was the manufactured cellulosics, 63.3 percent was the synthetics (67.9% manufactured), 31.1 percent was cotton, and less than 1 percent was wool. A significant increase was noted in the use of cotton during the eighties, from 3.0 billion lb in 1980 to 4.1 billion lb in 1990, believed to be in large measure due to advances in chemical finishes that made caring for the natural fibers easier than it was with the older technology. To some extent, this increase also may have occurred because people were willing to accept a more wrinkled look in order to gain the comfort of hydrophilic fibers. Still, the general figures represent quite a reversal in the consumption of fiber types seen three or more decades ago. For example, in 1960, of the 6.5 billion lb of fibers used, manufactured fibers accounted for only 29 percent of the total, with cotton 65 percent and wool 6 percent.

Major applications of fibers lie in apparel, home furnishing, and industrial products. In each of these, manufactured fibers have made large inroads, and currently their usage dominates. As an illustration, consider the changes that have taken place in the use of the materials required in the manufacture of tire cords. Originally made from cotton, rayon took a commanding position during World War II. But as late as 1951 cotton comprised about 40 percent of the total output

Fig. 21.3. Comparative growth of population and fiber consumption in the United States, 1960–90.

of tire cords of approximately half a billion lb, and nylon was at a negligible level of 4 million lb. By 1960, however, cotton had all but disappeared; nylon represented about 37 percent of the total (on a weight basis), even though only about 0.8 lb of nylon is needed to replace 1.0 lb of rayon. Whereas rayon for several years had dominated the so-called original-equipment tire market and nylon had held a corresponding position for replacement tires, more recently glass and polyester have made heavy inroads into both—especially in belted constructions. The situation continued to change in favor of noncellulosic manufactured fiber usage in tires, so that by 1972 rayon was down to 14 percent, nylon up to 42 percent, polyester up to 32 percent, glass up to 7 percent, and steel at 5 percent, all on a weight basis. By the late 1970s, tire markets were dominated wholly by manufactured fibers with polyester holding over 90 percent of the passenger car original-equipment market and nylon commanding over 90 percent of the truck original-equipment market. This division of markets is a direct result of the performance characteristics of the two fibers. Polyester-containing tires are free of "flat spotting" or cold-morning thump, and so are preferred in passenger cars for their smooth ride. On the other hand, nylon-containing tires are tougher and more durable, and so are the choice for trucks and off-road vehicles.

The production of manufactured fibers throughout the world has developed in a manner that rather parallels the situation in the United States, as may be seen in Fig. 21.4. There are some expected differences, and obviously the data for world usage are strongly influenced by the large components attributable to the United States, which currently accounts

Fig. 21.4. Production of fibers in the world, 1950–90.

for about 23 percent of the manufactured fiber and about 15 percent of the total fiber consumption. The output of the world cellulosics has leveled off, but expansion of the noncellulosics has continued unabated. The use, or at least the recorded use, of the natural fibers, cotton and wool, rose rapidly in the 1950s, as the world economy rebounded at the conclusion of World War II. Since then (1960–1970) a modest increase has continued, essentially parallel to the growth of world population. But, in comparison with population trends, it appears that the great demand has been for manufactured textile fibers. Much of this increase has resulted from an improved standard of living and the absence of major wars.

A detailed economic examination of the process of fibers and the changes that have taken place during the last half century would show two rather vivid occurrences. The first of these is a rapid decrease in the prices of the newer fibers as they became established, followed by a leveling out and stabilization. The second is the relative stability of prices of the manufactured fibers on short-term and even long-term bases, as compared with fluctuations in the prices for the natural fibers where governmentally imposed stability has not been in effect. Data are not presented about it in this text, but in the first half of the twentieth century there was a saying in the textile industry that the person who made or lost money for the company was the one who was responsible for buying cotton and wool "futures." However, it should also be emphasized that list prices of manufactured fibers are ceiling prices and do not reflect the short-term discounts, allowances, and special arrangements that are given in a free marketplace when the demand for any manufactured fiber softens.

A presentation of complete information about the consumption of raw materials, chemical reactions, reagents and catalysts used, and efficiencies of operation in the production of manufactured fibers undoubtedly would contribute to a better understanding

of the industrial chemistry involved. Several factors have prevented this, however. In the first half of the twentieth century, a historical belief in the efficacy of trade secrets still permeated the chemical industry. Even with the increased mobility of technical and scientific personnel during and following World War II, the idea still prevailed that if nothing other than patents was allowed to become public knowledge, so much the better. The situation has changed considerably since about 1960, as can be noted from the availability of information contained in the list of suggested readings that follows this chapter; yet, secrecy tends to be maintained despite the fact that key employees move from company to company, and the chemical engineering knowledge available in chemical companies that produce large volumes of fibers permits an almost complete appraisal of a competitor's activities.

In general, in the early period of production of a fiber, the cost of the original raw material may have had very little bearing on the selling price of the final fiber. A most important factor is the action of the producer's competitors and the conditions of the market and the demand that can be developed. But the complexity of the processes involved in conversion determines the base cost of the fiber at the point of manufacture. As the process becomes older, research reduces this complexity; with simplification, there may be rapid drops in plant cost. If demand remains high, such reductions will not be expected to be reflected in selling prices; rather, profits are high. As more producers enter the field in order to share in those profits, output capacity surpasses demand, and in accordance with classical economic theory, major selling price reductions result. This was happening, in general, in the 1960s and for cellulose-based manufactured fibers in the 1970s. But beginning in 1973, the cost of petroleum-based products started to rise steeply and erratically. This rise was based not on economic considerations alone but on political considerations among the oil producing and exporting countries (OPEC) as well. Further upward pressure on manufactured fiber prices has resulted from governmental limitations placed on chemical usage and exposure and on amounts of chemicals that can be discharged into the air and water. To meet these limitations, the manufactured fiber industry has had to supply large infusion of capital. In some instances, such expenditures could not be justified, and plant capacity was shut down permanently. This was particularly true in the case of filament rayon. In recent years, the factors of rapidly rising raw material/energy prices and the costs of meeting environmental regulations have not allowed the prices of manufactured fibers to fall as production experience has been gained and technological advances have been introduced. Instead, selling prices have been continually adjusted upward in an effort to pass along unavoidable cost increases so as to maintain profitability. In areas of application in which a manufactured fiber replaced a natural one because of lower prices and stable availability, swings in fashion and increases in imports sometimes have caused a reduction in fiber utilization. This scenario is believed to have particularly applied to acrylic fiber, whose production has decreased recently in both Europe and the United States.

The great importance of manufactured fibers in the chemical industry and in the overall economy of the United States becomes apparent when the volume of production of these materials is considered and compared with the market value of even the least expensive of the raw materials used by them. The amounts of oil and natural gas consumed by the manufactured fiber industry represent around one percent of national annual usage. Of this amount, about one-half is used to produce raw materials from petrochemicals, with the other half used for energy to convert trees to wood pulp for cellulose-based fibers and to convert the wood pulp and petrochemical-derived raw materials to fibers.

RAYON

Chemical Manufacture

Rayon, the first of the manufactured fibers produced in large volume, is based on the

natural polymer cellulose, a repeat unit of which is shown below:

Two anhydroglucose units

Although in the early days the main source of this raw material was cotton linters, a combination of improved technologies for obtaining alpha cellulose from wood and the shortage of cotton linters used for the manufacture of cellulose nitrate during World War II resulted in a shift in raw material to wood pulp. Only certain trees constitute the most economical supply of dissolving pulp, as the final product is called; the process economics depend upon the cost of logs delivered at the pulping mill and the relative yield of alpha cellulose after the unusable lignin and other components of the wood are discarded.

A general flow diagram for the manufacture of rayon is given in Fig. 21.5. The dissolving pulp is received by the rayon manufacturer in sheet or roll form. In the manufacturing process, impurities are removed, with special attention being given to removal of traces of such metallic elements as manganese and iron, the former having an effect on the manufacturing process (as will be noted later) and the latter an effect on the color of the final product. The production of dissolving pulp involves drastic chemical action at elevated temperatures, which substantially reduces the originally very high molecular weight of the cellulose. The portion not soluble in 17 to 18 percent aqueous caustic, known as alpha cellulose, remains, and the lower-molecular-weight beta and gamma fractions are largely soluble and lost. The composition of the pulp is aimed at high alpha content. A typical economic trade-off is involved. The pulp producers can secure an alpha content of up to 98 percent by means of a cold caustic extraction,

Fig. 21.5. Flow diagram for manufacture of viscose yarn: (1) cellulose sheets and caustic soda; (2) steeping press; (3) shredder; (4) xanthating churn; (5) dissolver; (6) caustic supply; (7) ripener; (8) filtration; (9) deaeration; (10) filtration; (11) continuous process; (12) tire cord; (13) pot spinning; (14) staple spinning.

or, on the other hand, the rayon manufacturer can use a less expensive, lower-alpha-content pulp (90–96%) and expect to secure a lower yield. The sellers have numerous grades available to meet the specific process needs and end-product requirements of each of the buyers.

In the manufacture of rayon it is the usual practice to begin "blending" at the first step, which involves steeping the pulp. Further blending proceeds throughout succeeding steps. The warehouse supply of pulp consists of numerous shipments, and in making up the batches for the conventional process, a few sheets are taken from each of several shipments.

This serves two purposes. It prevents a slight variation in a single pulp lot from unduly affecting any given volume of production, and it provides a moving average so that changes with time are reduced to a minimum.

The cellulose sheets are loaded vertically, but loosely, into a combination steeping bath and press (Fig. 21.6), which is slowly filled with a solution of 17 to 19 percent caustic, where they remain for about one hour. In the steeping, the alpha cellulose is converted into alkali or "soda" cellulose; at the same time, as already mentioned, the caustic solution removes most of the beta and gamma celluloses (also called hemicelluloses). The exact chemical composition of the soda cellulose is not known, but there is evidence that one molecule of NaOH is associated with two anhydro-glucose units in the polymer chain.

The excess caustic solution is drained off for reuse. Additional amounts are removed by forcing the sheets through a press. The sheets are still in a swollen state and retain from 2.7 to 3.0 parts of the alkali solution. The spent steeping solution squeezed out of the pulp is processed for recovering the caustic from the organic materials.

The sheets of soda cellulose are discharged into a shredder. If blending is desired, the charges from two or more steeping presses are mixed in a single shredder, where the already soft sheets are torn into crumbs; cooling is provided to prevent thermal degradation. Shredding is controlled to produce crumbs that are open and fluffy, and that will allow air to penetrate the mass readily; this is essential in aging.

$$(C_6H_{10}O_5)_n \text{ (cellulose)} + 18\% \text{ aqueous NaOH} \rightarrow [(C_6H_{10}O_5)_2 \cdot \text{NaOH}]_n \text{ (swollen, insoluble, soda cellulose I)} + \text{soluble soda cellulose from } \beta \text{ and } \gamma \text{ celluloses}$$

Fig. 21.6. Steeping of cellulose in the manufacture of viscose rayon. (*Courtesy Avtex Fibers, Inc.*)

Soda cellulose is aged by holding it at a constant temperature in perforated containers. The oxygen in the air produces uniform aging accompanied by a reduction in molecular weight and an increase in the number of carboxyl groups present. The target of aging is an average molecular weight high enough to produce satisfactory strength in the final fiber but low enough so that the viscosity of the solution will not be excessively high at the desired concentration for spinning. Each of the various rayon end products has its optimum degree of polymerization or chain length, ranging from about one-fourth the original length for regular rayon to one-half for certain high-performance fibers. As noted earlier, this optimum size is generally established by effecting a compromise between process economics and desired end-product properties. The aging proceeds for periods of up to two or three days, although the tendency is to speed up the operation by using higher temperatures and traces of metal ions such as manganese or cobalt to catalyze the reaction. A combination of experience and constant quality-control testing guarantees that the material will reach the correct point for conversion to cellulose xanthate.

Cellulose xanthate, or more exactly, sodium cellulose xanthate or sodium cellulose dithiocarbonate, is obtained by mixing the aged soda cellulose with carbon disulfide in a vapor-tight xanthating churn. Based upon weight of cellulose, the amount of carbon disulfide used will be in the range of 30 percent for regular rayon to 50 to 60 percent for modified varieties.

$$[(C_6H_{10}O_5)_2 \cdot NaOH]_n + CS_2 \text{ (30–60\% based on weight of cellulose in soda cellulose)} \rightarrow (C_6H_{10}O_5)_n[C_6H_7O_2(OH)_x(O{-}\underset{\underset{S}{\|}}{C}{-}S \cdot Na^+)_{3-x}]_m \text{ or,}$$

$$\text{for simplicity, (cellulose}{-}O{-}\underset{\underset{S}{\|}}{C}{-}SNa)$$

The xanthate is soluble in a dilute solution of sodium hydroxide—a characteristic discovered by Cross and Bevan in 1892—and this property makes the spinning of rayon possible. It is a yellow solid; when dissolved in a dilute solution of alkali, it becomes a viscous, honey-coloured liquid—hence the word "viscose." At this stage the viscous solution may contain about 7.25 percent cellulose as xanthate in about a 6.5 percent solution of sodium hydroxide, although concentrations of both vary, depending on what end products are desired. The solution is ready for mixing with other batches to promote uniformity, to be followed by filtration, ripening, deaeration, and spinning. The filtration process usually involves several stages so that filters of decreasing pore size may be used to secure a balance of throughput and stepwise particle and gel removal.

Such an operation is a straightforward one for "bright" rayon, but only in the days of "artificial silk" did the shiny fiber alone satisfy the market. After a few years, a dull-appearing fiber also was demanded. At first, fine droplets of oil in the filaments were used to produce dullness until it was discovered that titanium dioxide pigment having a particle size smaller than one micron in diameter was even more satisfactory. The latter has since become the universal delustrant for all manufactured fibers. With the use of pigments of any type, problems of dispersion and agglomerate formation must be faced. The usualy practice has been to add this pigment when mixing the cellulose xanthate into the dilute solution of caustic.

However, there are many other chemicals and additives that a producer may be required to add to the solution, including: (1) a few parts per million of a tracer element for later identification of the product; (2) coloring pigments for "dope dyed" rayon ("dope dyeing" will be discussed in greater detail under another heading); (3) chemicals for controlling the rate of precipitation and regeneration for obtaining rayon with so-called high performance; and (4) polymers and chemicals to

impart specific properties to the fiber. From the standpoint of chemical processing, it is obvious that these additives may also be added when the sodium cellulose xanthate is dissolved in dilute caustic solution, or may be injected into the solution before it enters the spinnerette prior to being extruded. To keep the operations as flexible as possible, the additives should be injected at the last possible moment so that when a changeover is desired there will be a minimum amount of equipment to be cleaned. On the other hand, the farther along in the operation that additives are placed into the stream, the greater the problem of obtaining uniformity in an extremely viscous medium, and the greater the difficulty in maintaining exact control of proportions before the viscous solution is passed forward and spun. Furthermore, all insoluble additives must be of extremely small particle size, and all injected slurries must be freed of agglomerates by prefiltration; if not, the viscous solution containing the additives must be filtered. Each manufacturer of viscose rayon develops the particular conditions for making additions, depending on a multitude of factors, not the least of which is the existing investment in equipment. All manufacturers must face the universal necessity of filtering the solution with or without pigments or other additives, so that all impurities and agglomerates that might block the tiny holes in the spinnerette are removed.

Although it was known in the years following the discovery by Cross and Bevan that a viscose type of solution could be used in the preparation of regenerated cellulose, the conversion of this solution into useful fibers was not possible until the discovery that the solution required aging until "ripe." Ripening is the first part of the actual chemical decomposition of cellulose xanthate, which, if allowed to proceed unhampered, would result in gelation of the viscose solution.

Experience has taught the manufacturer the correct time and conditions for the aging operation, but the requirement of aging itself demands that the entire process be so planned that the viscose solution will arrive at the spinnerette possessing, as nearly as possible, the optimum degree of ripeness, to produce fibers having the desired characteristics. This degree of ripeness is determined by an empirical test made periodically, which is a measurement of the resistance of the solution to precipitation of the soda cellulose when a salt solution is titrated into it. Thus it is known as the "salt index" or "Hottenroth number" after its originator. An additional step in the overall ripening operation involves the removal of dissolved and mechanically held air by the use of a vacuum on a moving thin film of the viscose solution.

It should be mentioned that so inevitable is decomposition of cellulose xanthate and consequent gelation of the contents of pipes and tanks, that all viscose rayon plants must be prepared to pump in-process viscose solutions to other spinning machines or to a waste receiver, purging the entire system with dilute caustic solution, in the event of a long delay in spinning.

Wet Spinning

Spinning a viscose solution into rayon fibers (wet spinning) is the oldest of the three common ways of making manufactured fibers. In this method, the polymer is dissolved in an appropriate solvent, and this solution is forced through fine holes in the face of the spinnerette, which is submerged in a bath of such composition that the polymer precipitates. The pressure necessary for this extrusion is supplied by a gear pump which also acts as a metering device; the solution is moved though a final or "candle" filter before it emerges from the holes of the spinnerette.

$$\text{Cellulose}-\text{O}-\underset{\underset{\text{S}}{\|}}{\text{C}}-\text{SNa} \xrightleftharpoons{\text{H}_2\text{O}} \text{Cellulose}-\text{O}\underset{\underset{\text{S}}{\|}}{\text{C}}-\text{SH} + \text{NaOH} \xrightleftharpoons{\text{H}_2\text{O}}$$

$$\text{Cellulose} + \text{HO}\underset{\underset{\text{S}}{\|}}{\text{C}}-\text{SH} \longrightarrow \text{Cellulose} + \text{CS}_2 + \text{H}_2\text{O}$$

There is immediate contact between these tiny streams and the liquid or "wet" bath. As the bath solution makes contact with the material extruded from the holes, chemical or physical changes take place. These changes, whether of lesser or greater complexity, convert the solution of high-molecular-weight linear polymer first to a gel structure and then to a fiber. As will be observed in what follows, it is an interesting fact that the spinning of viscose rayon, with all of the ramifications made possible by variations in the composition of the solution and the precipitating bath, as well as in the operating conditions, presents the chemist and the chemical engineer with both the oldest and the most complex wet-spinning process.

The formation of rayon fibers from viscose solution is far from being simple, from either a physical or a chemical standpoint. The spinning bath usually contains 1 to 5 percent zinc sulfate and 7 to 10 percent sulfuric acid, as well as a surface-active agent, without which minute deposits will form around the holes in the spinnerette. Sodium sulfate (15 to 22 percent) is present, formed by the reactions, and as sulfuric acid is depleted and sodium sulfate concentration builds up, an appropriate replenishment of the acid is required. There is a coagulation of the organic material as the sulfuric acid in the spinning bath neutralizes the sodium hydroxide in the viscose solution; at the same time, chemical decomposition of the sodium cellulose xanthate takes place to regenerate the cellulose. If zinc ions are present, which is the usual situation in the production of the improved types of rayon, an interchange takes place so that the zinc cellulose xanthate becomes an intermediate. It reacts at a slower rate, causing slower decomposition to cellulose. This provides conditions for more effective stretching or drawing of the fiber. Chemical additives usually are present to repress hydrogen ion action. The gel-like structure, the first state through which the material passes, is not capable of supporting itself outside the spinning bath. As it travels through the bath, however, it quickly becomes transformed into a fiber that can be drawn from the spinning bath and that can support itself in subsequent operation (Fig. 21.7). The reaction between the bath and the fiber that is forming are paramount in determining the characteristics of the final product; it is for this reason that additives (previously mentioned), as well as zinc ions, may be used to control both the rate of coagulation and regeneration. In this manner, the arrangement of the cellulose molecules may be controlled to produce the conformational structure desired. A practical application of this will be discussed later.

Fig. 21.7. Spinning of viscose rayon. (*Courtesy Avtex Fibers, Inc.*)

$$2\text{Cellulose}-\text{O}-\underset{\underset{\text{S}}{\|}}{\text{C}}-\text{SNa} + \text{H}_2\text{SO}_4 \xrightarrow{\text{(Rapid reaction)}} \text{Cellulose} + \text{Na}_2\text{SO}_4 + \text{CS}_2$$

$$2\text{Cellulose}-\text{O}-\underset{\underset{\text{S}}{\|}}{\text{C}}-\text{SNa} + \text{ZnSO}_4 \xrightarrow{\text{(Slow reaction)}} (\text{Cellulose O}-\underset{\underset{\text{S}}{\|}}{\text{C}}-\text{S})_2\,\text{Zn} + \text{Na}_2\text{SO}_4$$

$$(\text{Cellulose O}-\underset{\underset{\text{S}}{\|}}{\text{C}}-\text{S})_2\text{Zn} + \text{H}_2\text{SO}_4 \longrightarrow \text{Cellulose} + \text{ZnSO}_4 + \text{CS}_2$$

Because of hydraulic drag, stretching occurs in the bath and also in a separate step after the yarn leaves the bath. In both cases, the linear molecules of cellulose are oriented from random positions to positions more parallel to the fiber axis. If a rayon tire cord is to be the final product, the fibers must be severely stretched to produce a very high orientation of the molecule. This is the basis of the tire cord's high strength and ability to resist stretching, without which growth of the tire body would occur. For regular textile and nonwoven uses such high strengths are not desired, and the spinning and stretching conditions are controlled to produce rayon of lower strength and greater stretchability under stress.

In order to stretch the yarn uniformly during the manufacturing process, two sets of paired rollers or "godets" are employed, each of the two sets operating at different rotational speeds. The yarn is passed around the first set of godets several times to prevent slippage and is supplied to the stretching area at a constant speed. A second set of godets moves it forward at a more rapid rate, also without slippage. Stretching may range from a few to 100 or more percent. Spinning speeds are on the order of 100 meters/min, but may vary with both the size of the yarn and the process used.

Spinning conditions, composition of the spinning bath, and additives to the viscose solution determine the physical characteristics of the rayon—its breaking strength and elongation, modulus, ability to resist swelling in water, and characteristics in the wet state as compared with those of the dry material. Not only must the chemical composition of the spinning bath be carefully controlled, but the temperature must be regulated at a selected point, somewhere in the range of 35 to 60°C, to ensure those precipitation and regeneration conditions essential to the manufacture of any particular viscose rayon having the properties needed for a selected end use.

After precipitation and regeneration of cellulose have been completed and raw rayon fiber has been formed, the subsequent steps must be controlled so that differences in treatment are minimized; otherwise such sensitive properties as "dye acceptance" will be affected, and the appearance of the final product will vary.

Minute traces of suspended sulfur resulting from the chemical decomposition of cellulose xanthate must be removed by washing with a solution of sodium sulfide. It is expedient to bleach the newly formed fibers with hypochlorite to improve their whiteness; an "antichlor" follows. The chemicals originally present and those used to purify the fibers must be removed by washing. As a final step, a small amount of lubricant is placed on the filaments to reduce friction and improve processibility in subsequent operations.

Several different processes are used for the steps involved in spinning and purifying continuous filament rayon. One of the most common involves the formation of packages of yarn, each weighing several pounds, for separate treatment. After it has been passed upward out of the spinning bath and stretched to the desired degree, the yarn is fed downward vertically into a rapidly rotating canlike container called a spinning pot or "Topham" box (after the man who invented it in 1900). It is thrown outward to the wall of the pot by centrifugal force and gradually builds up like a cake, with excess water being removed by the same centrifugal force. This cake is firm, although it must be handled with care, and is sufficiently permeable to aqueous solution to permit purification.

In another method of package-spinning, the yarn is wound onto a mandrel from the side at a uniform peripheral speed. With this process, the yarn may be purified and dried in the package thus formed. in any of these systems, the spinning and stretching, as well as subsequent steps, may involve separate baths.

The continuous process for spinning and purifying textile-grade rayon yarn merits particular mention from the standpoint of industrial chemistry, as it is rather an axiom that a continuous process is to be preferred over a batch or discontinuous operation. This method employs "advancing rolls" or godets that make it possible for the yarn to dwell for a

suffcient length of time on each pair, thus allowing the several chemical operations to take place in a relatively small area. Their operation depends on the geometry existing when the shafts of a pair of adjacent cylindrical rolls are oriented slightly askew. Yarn led onto the end of one of these and then around the pair will progress toward the other end of the set with every pass, the rate of traversing, and therefore the number of wraps, being determined by the degree of skewness.

The production of rayon to be converted to staple fiber also is amenable to line operation. Here the spinnerette has many thousands of holes, and a correspondingly large number of filaments are formed in the precipitation bath. The resulting tow then is stretched to the desired degree and immediately cut in the wet and unpurified condition. The mass of short lengths can be conveyed through the usual chemical treatments, after which it is washed and dried. It is fluffed to prevent matting and is packaged for shipment in large cases.

Cuprammonium, Nitrocellulose, and Cellulose Acetate Processes for Rayon

Cuprammonium Cellulose. Cellulose forms a soluble complex with copper salts and ammonia. Thus when cellulose is added to an ammoniacal solution of copper sulfate that also contains sodium hydroxide, it dissolves to form a viscous blue solution, and in this form it is known as cuprammonium cellulose. The principles on which the chemical and spinning steps of this process are based are the same as those for the viscose process. Cellulose is dissolved, in this case in a solution containing ammonia, copper sulfate, and sodium hydroxide. Unlike the viscose solution, the cuprammonium solution need not be aged and will not precipitate spontaneously on standing except after long periods. It is, however, sensitive to light and oxygen. It is spun into water and given an acid wash to remove the last traces of ammonia and copper ions. Although this rayon was never manufactured in a volume even approaching that achieved by the viscose process, the smaller individual filaments inherent in it made it useful in certain specialty markets. It no longer is manufactured in the United States but continues to be made abroad.

Nitrocellulose and Cellulose Acetate. Although nitrocellulose and cellulose acetate intermediates have been made and regenerated to form cellulose fibers, neither of these historical processes is still in operation.

Textile Operations

After the filament rayon fiber has been spun and chemically purified, much of it passes through what are known as "textile operations" before it is ready to be knitted or woven. Because these steps of twisting and packaging or beaming are common to the manufacture of all manufactured fibers, it is advisable to review briefly the background and the processes.

Rayon, the first manufactured fiber, not only had to compete in an established field, but also had to break into a conservative industry. Silk was the only continuous filament yarn, and products made from it were expensive and possessed of high prestige, so that they offered a tempting market for rayon. Thus, the new product entered as a competitor to silk and, as already noted, became known as "artificial silk." Under the circumstances it was necessary for rayon to adapt itself to the then existing silk processing operations and technologies. It was customary to twist several silk filaments together to secure a yarn of the desired size, strength, and abrasion resistance. Because rayon was weaker than silk and its individual filaments were smaller, it required as much twisting as silk or even more.

This twisting could have been carried out in the same plant where the yarn was spun, but the existence of silk "throwsters" (from the Anglo-Saxon *thrawan*, to twist or revolve) made that unnecessary. However, as the rayon industry developed, the amount of yarn twisted in the producing plant or sent forward to throwsters decreased. Over the years the trend has been to use less twist and to place, instead, several thousand parallel ends directly on a "beam," to form packages weighing as much

as 300 to 400 lb, which are shipped directly to a weaving or knitting mill. The advent of stronger rayons, as well as other strong fibers, and a diminishing market for crepe fabrics which required highly twisted yarns, accelerated the trend away from twisting.

In all twisting and packaging operations, the yarn makes contact with guide surfaces and tensioning devices, often at very high speeds. To reduce friction it is necessary to add a lubricant as a protective coating for the filaments. This is generally true of all manufactured fibers, and it is customary to apply the lubricant or "spinning finish" or "spinning lubricant" as early in the manufacturing process as possible. For those materials that develop static charges in passing over surfaces, this lubricant also must provide antistatic characteristics.

It is difficult to overstate the importance of fiber lubricants to the successful utilization of manufactured fibers. Few problems can be more damaging to a fiber-handling operation than a lubricant upset. A separate chapter could be written on lubricant usage, but some of its more important aspects will be mentioned here. Obviously, lubricants must reduce friction between the fiber and various surfaces to allow movement without excessive damage to the fiber or the surface contacted, the latter being any one of a variety of metals or ceramics. Similarly, a fiber comes in contact with the surface of other fibers in staple fiber processing and in packaging. The lubricant composition must be stable under a variety of storage conditions, without decomposing or migrating within the package or being lost from the fiber surface by adsorption into the fiber, and must be nontoxic and nondermatitic. It also must be compatible with other materials added during textile processes, such as the protective size coat applied to warp yarns before weaving or the wax coat often applied to yarns before circular knitting. Possible metal corrosion must be evaluated for each lubricant composition. Finally, after having performed its function, the fiber finish or lubricant must be completely scoured from the fabric to permit uniform adsorption of dyes and fabric finishes. The application of sewing lubricants to fabric to be cut and sewn is yet another area requiring attention.

Spun Yarn. After rayon became established in the textile industry where it could be used as a silk-like fiber, and its selling price was greatly reduced, other markets for it were developed. The cotton, wool, worsted, and linen systems of converting short discontinuous fibers to yarns were well established, and their products were universally accepted. Here again it was necessary to make rayon fit the requirements of existing equipment and historically acceptable operations. The first of these was that it be cut into the same lengths as those found in cotton and wool. Fortunately, the viscose rayon process was and is eminently suited to the production of tows containing thousands of filaments. The pressure required to force the solution through the holes is so low that neither thick metal sections nor reinforcement of the surface is necessary to prevent bulging, and large spinnerettes containing several thousand holes can be used. Furthermore, the spinning bath succeeds in making contact with all the filaments uniformly. As a result, the spinning of viscose rayon tow is very similar in principle to the production of the smaller continuous filament yarns.

Because both cotton and wool possess distortions from a straight rodlike structure, machinery for their processing was designed to operate best with such crimped fibers. Thus, it was necessary for rayon staple to possess similar lengths and crimpiness in order to be adapted to existing equipment. The crimp, that is, several distortions from a straight path per inch, is produced in rayon "chemically" by modification of the structure. The precipitation-stretching step in spinning is carried out so that the skin and the core of the individual filaments are radially nonuniform and constantly changing over very short lengths along the filaments. Because the skin and the core differ in sensitivity to moisture, the two components shrink differentially, leading to the development of permanent distortions of the filaments. The latent chemical crimp may be enhanced by a thermo-mechanical step. In

this, the tow is fed between two wheels, which in turn force it into a chamber, called a "stuffer box," heated with steam, in which it is forced against the compacted material ahead of it, causing the straight filaments to collapse immediately. As the mass of material is pressed forward, it becomes tightly compacted, and it tends to remain axially distorted after it escapes through a pressure-loaded door at the opposite end of the chamber.

Modified Viscose Rayon Fibers

The variations of chemical ingredients, their concentrations, and the temperature of the spinning bath determine the rates of coagulation and regeneration, and thus the relative amounts of "skin" and "core" in the fiber's cross section. The skin is known to possess a higher degree of order and better mechanical properties than the core, so an increase in its proportion is desired in higher-performance fibers. The degree of orientation is determined by the stretch imparted. Inherent in process variations such as these is what may be called the "art" of viscose rayon manufacturing, whereby great diversity in rayon properties can be obtained. The cross-sectional morphologies of some rayons are illustrated in Fig. 21.8.[3]

High-Wet-Modulus Rayon. One of the important innovations in the rayon industry has been the development of High-Wet-Modulus (HWM) rayon. For its manufacture, the cellulose molecules require a higher degree of polymerization (DP) than regular rayon; so aging and ripening times are decreased, and the processing temperature is lowered. As the viscose solution flows through the spinnerette into the bath, coagulation takes place to form the needed skin. However, in order for an increase in tenacity to occur, regeneration should proceed slowly; this is aided by using a lower concentration of acid in the spin bath. Also, zinc is added because zinc ions in the bath slow down regeneration by forming zinc cellulose xathate, which is more resistant to acid decomposition than is sodium xanthate. Because of the slow generation process, what actually is formed is nearly an "all-skin" rayon with a round cross section. Because the structure also is stretched before it crystallizes, a higher stretch is possible; this gives a higher orientation. The fiber thus has significantly higher strength than regular rayon. This higher tenacity exists not only when the material is dry, but also when it is wet; hence the name high-wet-modulus rayon. Because of its high structural orientation and greater order, the fibers have fewer physically accessible sites for water molecules; thus it is less susceptible to swelling and to the adverse effects of basic cleaning solutions, so that the fiber's launderability is improved. The HWM rayons are used extensively for blending with cotton, wool, silk, and all other manufactured fibers.

High Absorbency Rayons. Over the past years, disposable products have become commonplace, especially in the United States and Europe. Cellulosic fibers, particularly rayons, have served the needs of the disposables industry because of their absorbent qualities. The most useful fibres for disposable/absorbent applications are the rayons with crenulations, crimp, and hollow regions, all of which add to the absorbency of the fiber. These characteristics are achieved in varying degrees by physical and chemical alterations in the spinning process. Crenulations, or random irregularities in the shape of the cross section, typical for most rayon fibers, are caused by the rapid formation of skin before the dehydration is complete. As the fiber interior loses solvent, it collapses in certain areas and produces the crenulated shape. Furthermore, fabricators have learned how to control the cross-sectional shape of filaments by using spinnerettes containing other than round holes. One example is fibers having Y-shaped cross sections, which have been found to be capable of picking up more water than those with round cross sections.

Hollow viscose fibers contain gas pockets produced by adding "blowing" agents, such as sodium carbonate, to the viscose. When carbon dioxide is released during regeneration, the fibers inflate, leading to the formation of hollow filaments. The added free volume and

Fig. 21.8. Cross-sectional morphologies of some of the rayon fibers. (*Source: Turbak, A., "Rayon," in* Encyclopedia of Polymer Science and Engineering, *2nd ed., Vol. 14, p. 55,* copyright © John Wiley and Sons, Inc. New York, 1985 and used with permission of the copyright owner.)

decrease in molecular order, increase the ability of the fibers to pick up water.

Other New Developments

A number of other developments are taking place in the rayon industry, the target of one being the manufacture of lint-free rayon for use in products such as circuit boards. In another, graphite particles are blended in with the viscose to reduce static buildup. Production of flame-retardant rayon is receiving increased attention, in one case being achieved by the addition of phosphorus compounds to the spinning dope. The advanced technology now can produce flame-resistant fibers that, when exposed to high temperatures, will not shrink or emit toxic gases. Other developments are

in the area of finding better and environmentally safer solvents for the cellulosic raw material. Searches are under way for solvents that may lead to lyotropic liquid crystalline polymer solutions from which ultra-high-strength and high-modulus fibers can be spun.

An example along the above lines may be the new solvent-spun fiber Tencel, by Courtaulds, which recently was introduced. It is claimed to have an environmentally clean process, with virtually total recycling of the solvent.[4] The manufacturing process is said to involve the dissolving of pulp in hot amine oxide, filtration of the solution, and then spinning into a bath containing a dilute solution of the solvent. The bath removes the amine oxide from the fibers, which are washed and dried, and the removed solvent is reclaimed for further use. The final fiber is said to have a different molecular structure from that of normal rayon, and a smooth surface and a round cross section. The fiber is claimed to be stronger than cotton and normal rayon in both the dry and the wet states.

CELLULOSE ACETATE

Historical

Cellulose acetate was known as a chemical compound long before its potential use as a plastic or fibre-forming material was recognized. The presence of hydroxyl groups had made it possible to prepare cellulose esters from various organic acids, as cellulose consists of a long molecular chain of beta-anhydro-glucose units, each of which carries three hydroxyl groups—one primary, the other two secondary. The formula for cellulose (already noted) is $[C_6H_7O_2(OH)_3]_n$; when this is fully esterified, a triester results. It was learned quite early that although cellulose triacetate is soluble only in chlorinated solvents, a product obtained by partial hydrolysis of the triester to a "secondary" ester (having about 2.35 to 2.40 acetyl groups per anhydro-glucose unit) was easily soluble in acetone containing a small amount of water. Many other cellulose esters have been prepared, but only the acetate has been commercialized successfully as a manufactured fiber. Propionates and butyrates, and mixed esters of one or both the acetate, have applications as plastics. The first acetate fibers were produced in 1921 in Europe and in 1924 in the United States.

Manufacture of Secondary Cellulose Acetate

Cellulose acetate originally was made from purified cotton linters, but this raw material has been entirely replaced by wood pulp. The other raw materials used are acetic acid and acetic anhydride.

Cellulose acetate is manufactured by a batch process (see Fig. 21.9). There has been mention in the patent literature of a continuous system, but its utilization as a production process has not been announced. The "charge" of cellulose, purified, bleached, and shredded, is of the order of 800 to 1500 lb. It is pretreated with about one-third its weight of acetic acid and a very necessary amount of water, about 6 percent of its weight. If it is too dry at the time of use, more H_2O must be added to the acetic acid. A small amount of sulfuric acid may be used to assist in swelling the cellulose and to make it "accessible" to the esterifying mixture.

Although there has been much discussion of the chemistry of cellulose acetylation, it is now generally agreed that the sulfuric acid is not a "catalyst" in the normal sense of the word, but rather that it reacts with the cellulose to form a sulfo ester. The acetic anhydride is the reactant that provides the acetate groups for esterification. The acetylation mixture consists of the output from the acetic anhydride recovery unit, being about 60 percent acetic acid and 40 percent acetic anhydride, in an amount 5 to 10 percent above the stoichiometric requirement, to which has been added 10 to 14 percent sulfuric acid based on the weight of cellulose used. The reaction is exothermic and requires that the heat be dissipated.

In preparing for acetylation, the liquid reactants are cooled to a point (0°C) where the acetic acid crystallizes, the heat of crystallization being removed by an appropriate

Fig. 21.9. Flow diagram for manufacture of cellulose acetate yarn: (1) wood pulp; (2) attrition mill; (3) cyclone; (4) 35 percent acetic acid; (5) pretreater; (6) magnesium acetate solution; (7) precooled acetylation mix; (8) sulfuric acid; (9) acetylator; (10) ripener; (11) steam; (12) blender; (13) precipitator; (14) dilute acetic acid; (15) hammer mill; (16) water; (17) rotary screen washer; (18) squeeze rolls; (19) drying oven; (20) blender; (21) storage bins; (22) silos; (23) weight bins; (24) acetone; (25) wood pulp; (26) pigment; (27) mixers; (28) hold tanks; (29) filter press; (30) pump; (31) filter; (32) air; (33) jet; (34) acetone recovery; (35) oiling wheel; (36) feed roll; (37) bobbin; (38) inspection.

cooling system. The slush of acetic acid crystals in the acetic anhydride–sulfuric acid mixture is pumped to the acetylizer, a brine-cooled mixer of heavy construction (see Fig. 21.10). The pretreated cellulose is dropped in from the pretreating unit located above. The reaction is highly exothermic, and at the start large amounts of heat are produced. As the temperature of the reaction mixture rises to the melting point of the acetic acid (16.6°C), its large heat of fusion (45.91 cal/g) prevents a dangerous rise in temperature that would degrade the molecular weight of the cellulose chain. As the reaction proceeds, brine in the jacket of the acetylizer provides additional cooling.

$$\text{Cellulose} + (CH_3CO)_2O + H_2SO_4(10\text{–}15\% \text{ based on weight of cellulose}) \xrightarrow[\text{Anhydrous}]{CH_3COOH} [C_6H_7O_2(OSO_3H)_{0.2}(CH_3COO)_{2.8}]_n$$

The reaction product is soluble in the acetylation mixture; as it is formed and dissolved, new surfaces of the cellulose are presented to the reagents. One variation of this procedure uses methylene chloride, rather than an excess of acetic acid in the reaction mixture. This chemical is used both to dissipate the heat by refluxing (boiling point, 41.2°C) and to dissolve the cellulose ester as it is formed. As the reaction proceeds, the temperature is allowed to rise. Because cellulose is a natural product obtained from many sources, it varies slightly in composition, and at the end of the reaction cannot be predicted exactly; the disappearance of fibers as determined by microscopic examination thus is the usual means of following its progress.

During the acetylation operation, a certain amount of chain fission is allowed to take place in the cellulose molecule. This is to

Fig. 21.10. Process vessel for acetylation of cellulose. (*Courtesy Celanese Fibers Co.*)

ensure that the viscosity of the cellulose acetate spinning solution will be low enough for ease of handling but high enough to produce fibers with the required strength. The temperature of the reaction controls the rates of both acetylation and degradation of molecular weight.

The next step in the manufacture is "ripening," whose object is to convert the triester, the "primary" cellulose acetate, to a "secondary" acetate having an average of about 2.35 to 2.40 acetyl and no sulfo groups (if any sulfuric acid is used in pretreatment) per anhydro-glucose unit. While the cellulose sulfo-acetate is still in the acetylizer, sufficient water is added to react with the excess anhydride and start the hydrolysis of the ester. Usually the water is used as a solution of sodium or magnesium acetate, which increases the pH and promotes hydrolysis. The temperature is raised to about 70 to 80°C, by direct injection of steam to speed the reaction. Hydrolysis is continued until the desired acetyl content is obtained. When this value is reached, an aqueous solution of magnesium or sodium acetate is added to cool the batch and stop the hydrolysis. It is then ready for precipitation. For example,

$$[C_6H_7O_2(OSO_3H)_{0.2}(CH_3COO)_{2.8}]_n + (CH_3COO)_2Mg \xrightarrow[\text{aqueous}]{\text{conc. } CH_3COOH} [C_6H_7O_2(OH)_{0.65}(CH_3COO)_{2.35}]_n + MgSO_4$$

The solution is carried to the verge of precipitation by adding dilute acetic acid. Then it is flooded with more dilute acetic acid and mixed vigorously, so that the cellulose acetate comes out as a "flake" rather than a gelatinous mass or fine powder. The flake then is washed by standard countercurrent methods to remove the last traces of acid, and is dried in a suitable dryer.

Manufacture of Cellulose Triacetate

To obtain completely acetylated cellulose, the reaction requires the use of perchloric acid rather than sulfuric acid as the catalyst. In the presence of 1 percent perchloric acid, a mixture of acetic acid and acetic anhydride converts a previously "pretreated" cellulose to triacetate without changing the morphology of the fibers. If methylene chloride rather than an excess of acetic acid is present in the acetylation mixture, a solution is obtained. However, usually a degree of substitution between 92 and 100 percent is acceptable. For obtaining such a triester, it is possible to use about 1 percent sulfuric acid instead of perchlolic acid. When the sulfoacetate obtained from such a reaction is hydrolyzed with the objective of removing only the sulfo-ester groups, the resulting product has about 2.94 acetyl groups per anhydroglucose unit. The preparation, hydrolysis, precipitation, and washing of "triacetate" are in all other respects similar to the corresponding steps in the manufacture of the more common secondary acetate. Cellulose triacetate, formerly produced under the trade name Arnel® by Celanese Corporation, is no longer in production in the USA.

Acid Recovery. In the manufacture of every pound of cellulose acetate about 4 lb of acetic

acid is produced in 30 to 35 percent aqueous solution. The accumulated acid contains a small amount of suspended fines and some dissolved cellulose esters. To remove the suspended material, the acid is passed slowly through settling tanks. Then it is mixed with organic solvents, so that the acid becomes concentrated in an organic layer, which is decanted. Distillation separates the acid from the organic solvent.

To produce the acetic anhydride, the acid is dehydrated to ketene and reacted with acetic acid using a phosphate catalyst at 500°C or higher in a tubular furnace.

$$CH_3COOH \xrightarrow[\text{catalyst}]{\text{heat}} H_2O + CH_2{=}C{=}O$$

$$CH_2{=}C{=}O + CH_3COOH \rightarrow (CH_3CO)_2O$$

The mixture of unreacted acid, water, and anhydride is fed to a still, which yields dilute acetic acid overhead and an anhydride–acetic acid mixture at the bottom (see Fig. 21.11). Conditions are controlled in such a way that the raffinate is about 40 percent anhydride and 60 percent acetic acid. As already mentioned, this is the desired ratio for the reaction mixture used for acetylation of cellulose.

Fig. 21.11. Recovery of acetic acid. (*Courtesy Celanese Fibers Co.*)

Blending of Flake. As in the manufacture of viscose, the products of batch operations are blended to promote uniformity in the manufacture of cellulose acetate. Although a blend of different celluloses is selected in the beginning, the pretreatment, acetylation, and ripening are batch operations with little or no mixing. Before precipitation, a holding tank provides an opportunity for mixing; then precipitation, washing, and drying—all continuous—promote uniformity. The dried cellulose acetate flake moves to holding bins for analysis—the moisture content, acetyl value, and viscosity being especially important. The results of the analyses determine what further blending is necessary. After blending and mixing of portions of selected batches, the lot is air-conveyed to large storage bins or "silos," which are filled from the center of the top and emptied from the center of the bottom, thus bringing about further mixing.

Spinning Cellulose Acetate

Acetone is metered into a vertical tank equipped with a stirrer, and the cellulose acetate flake and filter aid are weighed in an automatic hopper; all operations are controlled by proportioning methods common to the chemical industry. The ratio of materials is about 25 percent cellulose acetate, 4 percent water, less than 1 percent ground wood pulp as a filter aid, and the remainder acetone. The mixture moves forward through two or three stages at the rate at which it is used, the hold time being determined by experience. After dissolution is completed, filtration is carried out in batteries of plate and frame filter presses in three or even four stages, the passage of the "dope" being through presses of decreasing porosity.

Much of the cellulose acetate is delustered by the addition of titanium dioxide pigment, as with viscose rayon. Between filtrations (and after the last filtration), the dope goes to storage tanks that serve to remove bubbles; in this case, a vacuum is not necessary. From the final storage tank, it is pumped into a header located at the top of each spinning machine; then it is directed to a series of

metering gear pumps, one for each spinnerette. Because the holes in the cellulose acetate spinning spinnerette are smaller (0.03–0.05 mm) than those in the corresponding viscose devices, great care must be taken with the final filtration. An additional filter for the removal of any small particles that may have passed through the large filters is placed in the fixture, sometimes called the "candle," to which the spinnerette assembly is fastened. A final filter is placed in the spinnerette-assembly unit over the top of the spinnerette itself.

The method used for spinning cellulose acetate is "dry" spinning. The dope is heated (in some cases above the boiling point of acetone, 56.5°C) to lower its viscosity and thus reduce the pressure required to extrude it, and to supply some of the heat needed for evaporating the acetone solvent. (See Fig. 21.12.)

Fig. 21.12. Dry spinning of cellulose acetate. (*Courtesy Tennessee Eastman Co.*)

The spinnerette is stainless steel, and because the filaments must be heated and prevented from sticking together, and because space must be allowed for the escape of acetone vapor, the holes must be kept farther apart than those of the spinnerettes used for wet spinning. As the hot solution of cellulose acetate in acetone emerges downward into the spinning cabinet, an instantaneous loss of acetone takes place from the surface of the filaments, which tend to form a solid skin over the still liquid or plastic interior. A current of air, either in the direction the filaments are moving or countercurrent, heats the filaments, and as the acetone is diffused from the center through the more solid skin, each filament collapses to form the indented cross-sectional shape typical of cellulose acetate. The heated air removes the vaporized acetone. Each manufacturer uses a preferred updraft, downdraft, or mixed-draft operation, as needs dictate.

The cabinet through which the yarn passes vertically downward must be long enough to allow sufficient acetone to diffuse outward and evaporate from the surfaces of the filaments so that the latter will not stick to the first surface contacted or fuse to each other. The temperature of the air in the cabinet, the rate of flow, the length of the cabinet, the size and number of filaments, and the rate of travel are all interrelated in the spinning process. Because it is desirable to increase spinning speeds to the limit of the equipment, the tendency has been to construct longer spinning cabinets as each new plant is built. Present spinning speeds are on the order of 600 or more meters/min, measured as the yarn emerges from the cabinet.

Other dry-spinning operations have followed essentially the same pattern. For example, the dry spinning of cellulose triacetate was identical to that for secondary acetate except that the acetone solvent had been replaced by a chlorinated hydrocarbon such as methylene chloride, the solubility of which was improved by the addition of a small amount of methanol (5–15 percent).

The acetate yarn emerging from the cabinet makes contact with an applicator that provides

the lubricant required to reduce both friction and static formation in subsequent operations. With its surface lubricated, the yarn passes around a "feed" roll that determines the rate of withdrawal from the spinning cabinet, and then to any of several desired packaging devices.

Unlike the packaging of rayon yarn, cellulose acetates are either "ring" spun or wound into a package called a "disc," "zero twist," or "cam wound." In the ring-spun package, the yarn carries a slight twist of less than one turn per inch, but it requires a relatively expensive bobbin. Since the trend is toward less twisting, such acetate yarn is "beamed" in the producer's plant after little or no twisting, the heavy beams being shipped directly to knitters or weavers. (See Fig. 21.13.)

Filament yarns are twisted for two reasons. One is to supply certain aesthetic characteristics such as touch, drapability, and elasticity. The other more fundamental reason is to provide physical integrity to the filament bundle so that it can be warped, woven, and knitted without excessive breakage or fraying of individual filaments.

The yarn just mentioned as having no twist imparted before beaming may have been subjected to intermingling just prior to windup after extrusion. In the intermingling process, yarn with no twist, and usually under low tension, is passed through a zone where it is impinged upon by a jet stream of compressed air. This causes the filaments to interlace or intermingle with each other, and they can become metastable in this

Fig. 21.13. Beaming cellulose acetate yarn from a reel holding about 800 packages of yarn. (*Courtesy Tennessee Eastman Co.*)

configuration when tension is reapplied. In this condition, the yarn has the integrity of twisted yarn and will pass through several textile processing steps without difficulty; but with each handling, some of the intermingling is worked out.

Solvent Recovery. The air containing the acetone vapor is drawn out of the spinning cabinet and passed through beds of activated carbon that sorb the organic solvent. The acetone is recovered by steaming and then by separating it from the water by distillation. The efficiency of recovery is about 95 percent.

Dope-Dying. As with viscose rayon, colored pigments or dyestuffs may be added to the spinning solution so that the yarn will be colored as it is produced, thus eliminating the need for dyeing the final fabric. As mentioned earlier, even in using titanium dioxide, a compromise must be made on the basis of two competing needs. Complete mixing, uniformity, and filtration require that the addition be made early in the operation; minimal cleaning problems during changeovers require just the opposite. There exist two solutions to the problem. If a manufacturer must produce a multitude of colors in relatively small amounts, it is desirable to premix individual batches of spinning dope. Each batch should be pretested on a small scale to ensure that the desired color will be acceptable when it is produced. Facilities must be provided to allow each batch of colored dope to be cut into the system very close to the spinning operation in order to minimize pipe cleaning. Permanent piping must be flushed with solvent or the new batch of colored dope; some of the equipment may be disassembled for mechanical cleaning after each change of color.

Another method of producing spun-dyed yarn involves using a group of "master" dopes of such color versatility that when they are injected by appropriate proportioning pumps into a mixer located near the spinning operation, they will produce the final desired color. The advantages of such an operation are obvious; the disadvantage lies in the public demand for an infinite number of colors. No small group of known pigments will produce final colors of every desired shade.

PROTEIN FIBERS

As previously mentioned, the use of naturally existing polymers to produce fibers has had a long history. In the case of cellulose the results were fabulous. An initial investment of $930,000 produced net profits of $354,000,000 in 24 years for one rayon company.[5] On the other hand, efforts to use another family of natural polymers—proteins—have thus far resulted in failure or at best very limited production.

These regenerated proteins are obtained from milk (casein), soya beans, corn, and peanuts. More or less complex chemical separation and purification processes are required to isolate them from the parent materials. They may be dissolved in aqueous solutions of caustic, and wet-spun to form fibers, which usually require further chemical treatment as, for example, with formaldehyde. This reduces the tendency to swell or dissolve in subsequent wet-processing operations or final end uses. These fibers are characterized by a wool-like feel, low strength, and ease of dyeing. Nevertheless, for economic and other reasons they have not been able to compete successfully with either wool (after which they were modeled) or with other manufactured fibers.

NYLON

Historical

Nylon was the first direct product of the technological breakthrough achieved by W. H. Carothers of E. I. du Pont de Nemours & Co. Until he began his classic research on high polymers, the production of manufactured fibers was based almost completely on natural linear polymers. Such materials included rayon, cellulose acetate, and the proteins.

His research showed that chemicals of low molecular weight could be reacted to form polymers of high molecular weight. By selecting reactants that produce linear molecules having great length in comparison with their cross section, fiber-forming polymers are obtained. With this discovery the manufactured fiber industry entered a new and dramatic era.

Manufacture

Nylon 66. The word "nylon" was established as a generic name for polyamides, one class of the new high-molecular-weight linear polymers. The first of these, and the one still produced in the largest volume, was nylon 66 or polyhexamethylene adipamide. Numbers are used following the word "nylon" to indicate the number of carbon atoms contributed by the diamine and dicarboxylic acid constituents, in this case hexamethylenediamine and adipic acid, respectively.

To emphasize the fact that it does not depend on a naturally occurring polymer as a source of raw material, nylon often has been called a "truly synthetic fiber." To start the synthesis, benzene may be hydrogenated to cyclohexane:

$$C_6H_6 + 3H_2 \xrightarrow{\text{catalyst}} C_6H_{12}$$

or the cyclohexane may be obtained by fractionation of petroleum. The next step is oxidation to a cyclohexanol-cyclohexanone mixture by means of air:

$$xC_6H_{12} + O_2(\text{air}) \xrightarrow{\text{catalyst}} yC_6H_{11}OH + zC_6H_{10}O$$

In turn, this mixture is oxidized by nitric acid to adipic acid:

$$C_6H_{11}OH + C_6H_{10}O + HNO_3 \xrightarrow{\text{catalyst}} (CH_2)_4(COOH)_2$$

Adipic acid so obtained is both a reactant for the production of nylon and the raw-material source for hexamethylenediamine, the other reactant. The adipic acid first is converted to adiponitrile by ammonolysis and then to hexamethylenediamine by hydrogenation:

$$(CH_2)_4(COOH)_2 + 2NH_3 \xrightarrow{\text{catalyst}} (CH_2)_4(CN)_2 + 4H_2O$$

$$(CH_2)_4(CN)_2 + 4H_2 \xrightarrow{\text{catalyst}} (CH_2)_6(NH_2)_2$$

Another approach is through the series of compounds furfural, furane, cyclotetramethylene oxide, 1,4-dichlorobutane, and adiponitrile, as illustrated below. The furfural is obtained from oat hulls and corn cobs.

Or, 1,4-butadiene, obtained from petroleum, may be used as a starting raw material to make the adiponitrile via 1,4-dichloro-2-butene and 1,4-dicyano-2-butene:

$$CH_2{=}CHCH{=}CH_2 \rightarrow$$

$$ClCH_2CH{=}CHCH_2Cl \xrightarrow[\text{catalyst}]{\text{HCN}}$$

$$NCCH_2CH{=}CHCH_2CN \xrightarrow[\text{catalyst}]{H_2} NC(CH_2)_4CN$$

When hexamethylenediamine and adipic acid are mixed in solution in a one-to-one molar ratio, the "nylon salt" hexamethylenediammoniumadipate, the direct progenitor of the polymer, is precipitated. After purification, this nylon salt is polymerized to obtain a material of the desired molecular weight. It is heated to about 280°C under vacuum while being stirred in an autoclave for 2 to 3 hours; a shorter holding period follows; and the process is finished off at 300°C. The molecular weight must be raised to a level high enough to provide a fiber-forming material, yet no higher. If it is too high, the corresponding

Fig. 21.14. Flow diagram for the manufacture of nylon 66 yarn: (1) air; (2) cyclohexane from petroleum; (3) reactor; (4) recycle cyclohexane; (5) still; (6) cyclohexanol–cyclohexanone; (7) nitric acid; (8) converter; (9) adipic acid solution; (10) still; (11) impurities; (12) crystallizer; (13) centrifuge; (14) impurities; (15) adipic acid crystals; (16) dryer; (17) vaporizer; (18) ammonia; (19) converter; (20) crude adiponitrile; (21) still; (22) impurities; (23) hydrogen; (24) converter; (25) crude diamine; (26) still; (27) impurities; (28) nylon salt solution; (29) reactor; (30) stabilizer; (31) calandria; (32) evaporator; (33) excess water; (34) autoclave; (35) delustrant; (36) water sprays; (37) casting wheel; (38) polymer ribbon; (39) grinder; (40) polymer flake; (41) spinning machine; (42) heating cells; (43) spinnerette; (44) air; (45) drawtwisting; (46) inspection; (47) nylon bobbin. (*Note*: Whenever the demand for liquid polymer at a spinnerette is large, as, for example, in the spinning of tire yarn, it is pumped directly from the autoclave.)

viscosity in the subsequent spinning operation will require extremely high temperatures and pressure to make it flow. Accordingly, a small amount of acetic acid is added to terminate the growth of the long-chain molecules by reaction with the end amino groups.

The polymerized product is an extremely insoluble material and must be melt-spun, as discussed later. Therefore, should a delustered or precolored fiber be desired, it is necessary to add the titanium dioxide or colored pigment to the polymerization batch prior to solidification. For ease of handling, the batch of nylon polymer may be extruded from the autoclave to form a thin ribbon, which is easily broken down into chips after rapid cooling. But, whenever possible, the liquid polymer is pumped directly to the fiber melt spinning operation. (See Fig. 21.14.)

Nylon 6. Nylon 6 is made from caprolactam and is known as Perlon® in Germany, where it was originally developed by Dr. Paul Schlack.[6] Its production has reached a very large volume in the United States in recent years.

Like nylon 66, nylon 6 uses benzene as raw material, which is converted through previously mentioned steps to cyclohexanone. This compound is in turn converted to the corresponding oxime by reaction with hydroxylamine, and cyclohexanone oxime is made into caprolactam by the Beckmann rearrangement.

After purification, the lactam is polymerized by heating at elevated temperatures in an inert atmosphere. During self-condensation, the ring structure of the lactam is opened so that the monomer acts as an epsilon-aminocaproic acid radical. Unlike that of nylon 66, the polymerization of caprolactam is reversible; the polymer remains in equilibrium with a small amount of monomer. As with nylon 66, nylon 6 is extruded in thin strands, quenched, and cut into chips for subsequent spinning,

or the molten polymer is pumped directly to the spinning equipment.

Melt spinning

Because of its extremely low solubility in low-boiling and inexpensive organic solvents, nylon 66 required a new technique for converting the solid polymer into fibers; hence the development of "melt" spinning, the third basic method for producing manufactured fibers. The following description refers essentially to nylon 66 because it was the first to use the method, but the process applies, in general, to all melt-spun manufactured fibers.

In the original production of nylon fiber by melt spinning, the chips of predried polymer were fed from a chamber onto a melting grid whose holes were so small that only passage of molten polymer was possible. Both solid and liquid were prevented from contacting oxygen by maintaining an inert nitrogen atmosphere over the polymer supply. The polymer melted in contact with the hot grid and dripped into a pool where it became the supply for the spinning itself. This melting operation has been entirely replaced by delivery of the molten polymer pumped directly from the polymerization stage or by "screw" melting. In the latter process, the solid polymer in chip form is fed into an extrusion-type screw contained in a heated tube. The depth and the helix angle of the grooves are engineered in such a way that melting takes place in the rear section, and the molten polymer is moved forward under increasing pressure to a uniformly heated chamber preceding the metering pump.

Whatever means is used to secure the molten polymer, it is moved forward to a gear-type pump that provides both high pressure and a constant rate of flow to the final filter and spinnerette. The filter consists of either sintered metal candle filters, several metal screens of increasing fineness, or graded sand arranged in such a way that the finest sand is at the bottom. After being filtered, the molten polymer at a pressure of several thousand pounds per square inch is extruded through the small capillaries in the heavily constructed spinnerette. It is necessary to maintain the temperature of the pool, pump, filter, spinnerette assembly, and spinnerette at about 20 to 30°C above the melting point of the nylon, which is about 264°C for nylon 66 and 220°C for nylon 6. Fibers having desired cross sectional shapes can be produced by selecting spinnerettes containing holes of appropriate configuration. An example of a trilobal spinnerette capillary and the shape of the resulting trilobal fiber is given in Fig. 21.15.

The nylon production process requires that the extruded fibers emerge from the spinnerette face into a quench chamber where a cross current of relatively cool quench air is provided to promote rapid solidification. The solid filaments then travel down a chimney

Fig. 21.15. Scanning electron micrographs of: (top) trilobal shaped spinnerette, and (bottom) resulting cross sections of nylon fibers. The two pictures were taken at different magnifications. (*Courtesy BASF Corporation.*)

to cool further, and a lubricant is applied before they make contact with the windup rolls in order to prevent static formation and to reduce friction in subsequent textile operations. The freshly spun yarn from the spinning chamber is taken up by a traversing winder onto a yarn package and "drawn" in a separate operation. In modern high-speed processes, drawing still is required, but in many cases this is combined with spinning in a single operation, as will be described in what follows.

Drawing

It was learned early that the "as-spun" fibers made from nylon 66 could be extended to about four times their original length with very little effort, but that thereafter a marked resistance to extension took place. It was discovered that during this high extension the entire length of fiber under stress did not extend uniformly. Rather, a "necking down" occurred at one or more points, and when the entire length under tension had passed through this phenomenon, a high-strength fiber was obtained. It also was found that when more than one necking down was allowed to take place in a given length of fiber, a discontinuity occurred at the point where the two came together. Accordingly, the drawing operation was aimed at forcing the drawing to occur at a single point as the yarn advanced from the supply to the takeup package.

Where still used, cold drawing consists essentially of removing the yarn from the package prepared in the melt spinning operation and feeding it forward at a uniformly controlled rate under low tension. It is passed around a godet or roller that determines the supply rate and prevents slippage; for nylon 66 it then is wrapped several times around a stationary snubbing pin. From there it goes to a second roller that rotates faster than the supply roller to produce the desired amount of stretch, usually about 400 percent. The necking down occurs at the pin. In the case of nylon 6, drawing may be effected satisfactorily without passing the yarn around such a snubbing pin.

The long molecules of the nylon 66 or 6 polymer, which are randomly positioned in the molten polymer, when extruded from the spinnerette tend to form "crystalline" areas of molecular dimensions as the polymers solidify in the form of freshly spun fibers. In the drawing operation, both these more ordered portions as well as the amorphous areas tend to become oriented so that the lengthwise dimensions of the molecules become parallel to the long axis of the fiber, and additional intermolecular hydrogen bonding is facilitated. It is this orientation that converts the fiber having low resistance to stress into one of high strength.

By controlling the amount of drawing as well as the conditions under which this operation takes place, it is possible to vary the amount of orientation and the degree of crystallization. A minimal amount is preferable in the manufacture of yarns intended for textile applications wherein elongation of considerable magnitude and low modulus or stiffness is required rather than high strength. On the other hand, strength and high modulus are at a premium when fibers are to be used in tire cords and other industrial applications. High resistance to elongation is imperative if the tire is not to grow under conditions of use. In this connection it should be noted that nylon tire cord that has been produced by twisting the original tire yarn and plying the ends of these twisted yarns together is hot-stretched just before use at the tire plant to increase strength and reduce even further the tendency to elongate under tension.

The separate operations of spinning and drawing nylon presented a challenge whose object was combination of the two operations into a single continuous step. But the problem was obvious, for the operating speeds of the two separate steps already had been pushed as high as was thought to be possible. How then would it be possible to combine them into a continuous spin-draw, wherein a stretching of about 400 percent could take place? The answer lay in the manner in which the cooling air was used and in the development of improved high-speed winding devices. By first cooling the emerging fibers by a

concurrent flow of air and then cooling them further by a countercurrent flow, the vertical length of the cooling columns can be kept within reason. In-line drawing may occur in one or two stages, and relaxation may be induced if needed. The final yarn is said to be packaged at speeds of 6000 meters/min.

Other Nylons, Modifications, and New Developments

Although nylon 66 and 6 account for most of the polyamide fibers produced, a great many others have been experimentally synthesized and have been developed and manufactured in commercial amounts. Of these, some have been made into fibers, some with limited economic success. These nylons are identified by either the same numbering system used for nylon 66 or 6 or by a combination of numbers and letters, as follows:

Nylon 3	$\text{-}\!\!\left(\text{NH}\text{—}(\text{CH}_2)_2\text{CO}\right)\!\!\text{-}_n$
Dimethyl nylon 3	$\text{-}\!\!\left(\text{NH}\text{—}\text{C}(\text{CH}_3)_2\text{—}\text{CH}_2\text{—}\text{CO}\right)\!\!\text{-}_n$
Nylon 4	$\text{-}\!\!\left(\text{NH}\text{—}(\text{CH}_2)_3\text{—}\text{CO}\right)\!\!\text{-}_n$
Nylon 6T	$\text{-}\!\!\left(\text{NH}\text{—}(\text{CH}_2)_6\text{—}\text{NH}\text{—}\text{CO}\text{—}\langle\bigcirc\rangle\text{—}\text{CO}\right)\!\!\text{-}_n$
Nylon 7	$\text{-}\!\!\left(\text{NH}\text{—}(\text{CH}_2)_6\text{—}\text{CO}\right)\!\!\text{-}_n$
Nylon 12	$\text{-}\!\!\left(\text{NH}\text{—}(\text{CH}_2)_{11}\text{—}\text{CO}\right)\!\!\text{-}_n$
Nylon PACM-12	$\text{-}\!\!\left(\text{NH}\text{—}\langle\bigcirc\rangle\text{—}\text{CH}_2\text{—}\langle\bigcirc\rangle\text{—}\text{NH}\text{—}\text{CO}\text{—}(\text{CH}_2)_{10}\text{—}\text{CO}\right)\!\!\text{-}_n$
Nylon 46	$\text{-}\!\!\left(\text{NH}\text{—}(\text{CH}_2)_4\text{—}\text{NH}\text{—}\text{CO}\text{—}(\text{CH}_2)_4\text{—}\text{CO}\right)\!\!\text{-}_n$
Nylon 610	$\text{-}\!\!\left(\text{NH}\text{—}(\text{CH}_2)_6\text{—}\text{NH}\text{—}\text{CO}\text{—}(\text{CH}_2)_8\text{—}\text{CO}\right)\!\!\text{-}_n$

Dimethyl nylon 3 is solution-spun because it tends to decompose during melt spinning. Nylon 4 has a moisture regain (mass water per unit mass of dry fiber, under standard atmospheric conditions of 20°C and 65% RH) of 6 to 9 percent[7] and therefore is superior to other nylons for textile usages, being comparable to cotton. Nylon 11 was developed in France and has been trademarked as Rilsan®. It has a moisture regain of 1.8 percent and density of 1.04 g/cc as compared with 4 percent and 1.14 g/cc, respectively, for nylon 66. Nylon 7 is made in the former Soviet Union and marketed under the name Enant. The fiber has better stability to heat and ultraviolet light than nylon 66 and 6. Nylon 6T, an aromatic polymer, has a much higher melting point (370°C), a higher density (1.21 g/cc), and slightly higher moisture regain (4.5%) than nylon 66. It also has superior resistance to nylon 66 against heat. This fiber has served as a precursor to the development of aramid fibers. Nylon PACM-12, formerly produced under the trade name Qiana® in the United States, is no longer in production.

A *Chemical and Engineering News* report[8] suggests that the most serious competition to nylon 66 and 6 will be provided by a new, still experimental fiber, nylon 46, being developed by DSM in the Netherlands. Trade-named Stanyl®, this fiber results from the interaction of 1,4-diaminobutane and adipic acid. Better order in the structure in the fiber leads to greater crystallinity and, thus, to greater density (1.18 g/cc). The fiber has a melting point of about 300°C, and a breaking stress or tenacity of 9.5 gram force/denier ($\sim$1 GPa), modulus at 120°C of 20 gram force/denier (2.1 GPa), and shrinkage in gas at 160°C of 3 percent.

Some of the outstanding characteristics of nylon that are responsible for its many uses in apparel, home furnishing, and industrial products are its high strength and toughness, elastic recovery, resilience, abrasion resistance, and low density. Among many applications of the fiber are such products as intimate apparel and foundation garments, sportswear,

carpets, parachutes, tents, sleeping bags, and tire cords.

The world production of nylon has doubled in recent years, increasing from 3.8 billion lb in 1970 to about 7.5 billion lb in 1990. The fiber accounts for about 24 percent of the synthetic fibers produced worldwide. In the United States, the production of the fiber also doubled in two decades, increasing from 1.2 billion lb in 1970 to about 2.4 billion lb in 1990.

POLYESTERS

Historical

The stimulus for the development of polyester, as for nylon, was provided by the fundamental work of Carothers. Although his team's initial work was directed toward this material, because of greater promise shown by polyamides at the time the developmental work on polyesters was temporarily set aside. The polymer, however, attracted interest in Great Britain, where J. T. Dickson and J. R. Whinfield experimented with it and developed a successful polyester fiber.[5] They found that a synthetic linear polymer could be produced by condensing ethylene glycol with terephthalic acid or by an ester-exchange between the glycol and pure dimethyl terephthalate. The polymer thus obtained could be converted to fibers having valuable properties, including the absence of color. Like nylon, this material has been popularized under its generic name, polyester or just "poly." Those persons working with it commonly refer to it as PET. It first appeared under the trade name Terylene® (Imperial Chemical Industries, Ltd.,) in England, and was first commercialized in the United States in 1953 as Dacron® (E. I. du Pont de Nemours & Co).

Manufacture

When the development of polyethylene terephthalate (PET) occurred, ethylene glycol already was being produced in large amounts from ethylene, a by-product of petroleum cracking, by the oxidation of ethylene to ethylene oxide and subsequent hydration to ethylene glycol, which, in a noncatalytic process, uses high pressure and temperature in the presence of excess water.

On the other hand, although *o*-phthalic acid, or rather its anhydride, had long been produced in enormous amounts for use in the manufacture of alkyd resins, the *para* derivative was less well known and not available on a large scale. The synthesis is a straightforward one, however, from *p*-xylene, which is oxidized to terephthalic acid, either by means of nitric acid in the older process or by air (catalyzed) in the newer one. In the early years this compound then was converted to the easily purified dimethyl ester in order to obtain a colorless polymer adequate for the manufacture of commercially acceptable fibers.

Several other methods were developed for producing the desired dimethyl tetephthalate. The Witten (Hercules) process goes from *p*-xylene to toluic acid by oxidation of one of the methyl groups on the ring, following which the carboxyl group is esterified with methanol. This process then is repeated with the second methyl group to secure the dimethyl ester of terephthalic acid.

Either phthalic anhydride or toluene, both in ample supply as raw materials, can be used in the Henkel processes. Use of phthalic anhydride depends only upon dry isomerization of the potassium salt of the *ortho* derivative to the *para* form at about 430°C and 20 atmospheres pressure; or toluene is oxidized to benzoic acid, whose potassium salt can be converted to benzene and the potassium salt of terephthalic acid by disporportionation.

The first step in the reaction of dimethylterephthalate and ethylene glycol is transesterification to form bis(*p*-hydroxyethyl) terephthalate (bis-HET) and eliminate methanol.

This product then is polymerized in the presence of a catalyst to a low molecular weight compound and the by-product glycol is eliminated. In a second stage, at a temperature of about 275°C and under a high vacuum, the molecular weight is raised to secure the melt viscosity desired for the particular material involved. Like nylon, this final material may be extruded, cooled, and cut into chips for storage and remelting, or it may be pumped directly to the spinning machines.

From the beginning, it was obvious that there would be considerable progress in industrial chemistry, to say nothing of cost reduction, if the process could be simplified by making it unnecessary to go through the dimethyl derivative to secure a product of adequate purity. This was accomplished in the early 1960s when methods of purifying the crude terephthalic acid were developed, and conditions and catalysts were found that made possible the continuous production of a color-free polymer. It is said that the selection of the catalyst is especially aimed at the prevention of ether linkages in the polymer chain due to intracondensation of the glycol end groups.

Two additional rather similar routes are known. Both depend upon the reaction between ethylene oxide, rather than ethylene glycol, and terephthalic acid to form the bis-HET monomer already mentioned. The difference between the two methods lies in the point where purification is done: in one case, it is the crude terephthalic acid; in the other, it is the bis-HET monomer. In both cases this monomer is polymerized by known procedures to form a fiber-grade polyester. The titanium dioxide delustrant is added, as might be expected, early in the polymerizing process.

Another polyester that has reached long term commercialization is now produced in limited volume as Kodel 200® by Tennessee Eastman Co. and is considered to be 1–4 cyclo hexylene dimethylene terephthalate. The glycol that is used instead of ethylene glycol in this process exists in two isomeric forms, one melting at 43°C and the other at 67°C. This makes possible their separation by crystallization, to secure the desired ratio of the two forms for conversion to the polymer. This ratio determines the melting point of the polymer, a most important property for a material that is to be melt-spun. The polymer from the 100 percent *cis* form melts at 275°C, and that from the 100 percent *trans* form at 318°C. Indications are that the commercial product is about 30/70 *cis–trans*.

In 1973, the Federal Trade Commission modified the generic definition of polyester to include in the polyester category materials that previously were polyester ethers or benzoate polyesters. As a result, the fiber known as poly(ethylene oxybenzoate) or PEB and manufactured under the trade name A-Tell in Japan came to be known as polyester. This material is made by reacting parahydroxybenzoic acid and ethylene oxide

to give paraoxyethylenebenzoic acid, which is then polymerized to obtain polyethylenebenzoate:

The fiber softens at about 200°C and melts at 225°C. It is said to have a silk-like hand and appearance and other properties comparable to those of other polyesters.

Polyesters are melt-spun in equipment essentially the same as that used for nylon, already described. Wherever the volume is large and the stability of demand is adequate, the molten polymer is pumped directly from the final polymerization stage to the melt-spinning machine. The molten polymer is both metered and moved forward at high pressure by use of an extruder coupled with a gear-type pump, through filters to the spinnerette, which contains capillaries of about 9 mils (230 microns) diameter. Great care is used to eliminate moisture and oxygen from the chips, if they are used, and from the spinning chamber. When the polyester fibers are destined to become staple, the emerging filaments from a number of spinnerettes are combined to form a tow, which can be further processed as a unit. Continuous filament yarn is packaged for further processing such as drawing or texturing. Spin-drawing, described later, has become commonplace today and represents major cost savings to the fiber manufacturer.

Drawing

Unlike nylon, which in the as-spun state contains a high amount of crystalline component, polyethylene terephthalate fibers are essentially amorphous as spun. In order to secure a usable textile yarn or staple fiber, this product must be drawn under conditions that will result in an increase in both molecular orientation and crystallinity. This is done by drawing at a temperature well above the glass transition point, T_g, which is about 80°C. Conditions of rate and temperature must be selected so that the amorphous areas are oriented, and crystallization will take place as the temperature of the drawn fibers drops to room temperature. An appropriate contact-type hot plate or other device is used, and about 300 to 400 percent extension is effected. As with nylon, the conditions of draw, especially the amount, determine the force–elongation properties of the product. Industrial-type yarns, such as those intended to be used as tire cord, are more highly drawn than other yarns and have higher strength with less elongation. Figure 21.16 shows a drawn single filament.

Fig. 21.16. Drawnecking in polyester single filament. (*Courtesy E. I. du Pont de Nemours and Co.*).

Heat Setting

The ability of textile fibers to be "set" is not characteristic of manufactured fibers alone. Aided in many cases by the presence of starch, cotton fabrics can be ironed to a smooth and wrinkle-free condition; also, a sharp crease in wool trousers has been commonplace for generations. In other words, these fabrics were exposed to moisture at elevated temperatures while being held or pressed into desired geometrical configurations and then allowed to cool before being released from constraint. Such fabrics tend to remain unchanged while

Fig. 21.17. Skin-peeling in polyester filament showing fibrillar structure. (*Courtesy E. I. du Pont de Nemours and Co.*)

cool and dry, even though the fibers from which they are formed carry internal stresses; but reversion takes place upon washing or exposure to high relative humidity.

With the development of nylon, and especially polyesters, a durable kind of setting has become possible. When fabrics made from these fibers (see Fig. 21.17) are shaped and then exposed to elevated temperatures either in the dry condition or, in the case of nylon particularly, in the presence of water vapor, thermoplastic relaxation of induced stresses in the fiber takes place and configurations at the molecular level adjust to a new and lower energy level. This depends on not only the temperature used but also the duration of the exposure. Thus a few seconds at 230°C will produce the same results as exposure for a considerably longer period at a temperature 50 to 75 degrees lower. The permanency of the setting, that is the ability of a fabric or garment to return to its original configuration after temporary distortion even while exposed to moisture and raised temperatures, is a function of the severity of the heat setting. To impart true permanence, it is essential that the internal crystalline structure be annealed.

It is this property of polyamides and polyesters that has been the main factor contributing to "ease of care" and the "wash and wear" characteristics of garments made from these polymers. In turn, these garments have revolutionized both the textile and the apparel industries.

Textured Yarns

Fundamentally, the manufacture of "textured" yarns is closely related to the heat setting of fabrics, which must be composed of thermoplastic fibers such as nylon or polyester, the difference being that the individual filaments or bundle of filaments in textured yarns are distorted from an essentially straight rodlike form and then heat-set. In some instances, the fibers are distorted in a more or less random way; at other times a regular pattern is introduced.

The first commercially successful textured yarn was produced by highly twisting nylon 66, heat-setting it as a full package of yarn, and then untwisting it through zero and a small amount of twist in the opposite direction. This process changed yarn from a close-packed structure to one that was voluminous because of mutual interference of distorted filaments. The technique of heat-setting the twisted yarn as a batch-unit operation now has been replaced by a continuous operation, using what is known as a "false twisting" process. This is based upon the principle that if a length of yarn is prevented from rotating at both ends but is rotated on its axis at its center point, the resulting two sections will contain both "Z" and "S" twists in equal amounts. When this occurs with a moving yarn, any element in it will first receive a twist in one direction, but after passing the false twisting point must revert to zero twist. If it is then made to pass over a hot plate while in the twisted state and is heat-set in that configuration, even after returning to the untwisted condition the individual filaments will tend to remain distorted when lengthwise stress is released. Because of the low mass and diameters of textile yarns or monofilaments, it is possible to false-twist them at extremely high rotational speeds. Yarn forward speeds of about 1000

meters/min are currently obtainable by passing the yarn between, and in contact with, high-speed-friction twisting discs. (When attempts are made to secure higher rates, problems of twist control develop.) The same technique is now more commonly applied to unoriented (undrawn) or partially oriented yarn (POY) at the draw-texturing machine. The resulting yarn may be heat-set as part of the same continuous operation by passing it through a second heater under conditions of overfeed or little or no tension in order to secure both thermally stable geometric configurations in the individual distorted filaments that comprise the yarn, and the degree of "stretchiness" and bulk desired in the final product.

Because these yarns are being made in one less step and also within the plants spinning the parent product, this latest development may be said to constitute another advance in the industrial chemical technology of manufactured-textile products. This draw-texturing appears to be especially applicable to polyester yarns intended for fabrics known as "double knits" and "textured wovens."

Yarn can be forced forward by means of "nip" rolls, although this may seem to be quite contrary to the old adage that one cannot push on an end of string. When this is done so that the yarn is jammed into a receiver (stuffer-box) already full of the preceding materials, it collapses with sharp bends between very short lengths of straight sections. In this condition heat is applied, usually in the form of superheated steam, to set it. In practice, the mass of such yarn is pushed through a heated tube until it escapes at the exit past a spring-loaded gate. During this passage it is heat-set in a highly crimped configuration; then it is cooled before being straightened and wound onto a package. In another continuous process, the yarn or monofilament is pulled under tension over a hot sharp edge so that it is bent beyond its elastic limit and is heat-set in that condition. The process is known as "edge crimping," and the result is not unlike that produced by drawing a human hair over the thumbnail. The process is not used much today, but a yarn with similar crimp is produced by bicomponent spinning.

When such yarns are knitted or woven into fabric, the filaments tend to return to the configurations in which they were originally heat-set. Contraction takes place in the direction of the yarn axis, and this in turn converts the smooth flat fabric into a "stretch" fabric and gives the surface a textured appearance. These fabrics or the garments made from them, whatever the process used to produce the yarns, may be given additional

Fig. 21.18. Schematic comparison of the shapes of some of the textured yarns.

heat treatment to secure stability in a desired geometric configuration. A degree of stretch may be retained, or a flat and stable textured surface may be produced. There are a number of variations of the texturing process, which, combined with the many possibilities of heat setting, impart considerable versatility to the final product (see Fig. 21.18). The growth in the use of these products in the 1960s is well known. Carpeting also provides a significant market for them, as texture is one of the most important characteristics of soft floor coverings. Such products have been important to the successful use and expanded development of nylon and polyester yarns.

Staple Process

Unlike nylon, which is used mostly in the form of continuous filament yarn, polyester is utilized both in staple and in continuous filament form in large volumes. For producing staple fiber, the spinning machines consist of a series of packs, 10 to 30, each with 1000 to 3000 holes. The extrudates from different packs after solidification and application of finish are combined to form a subtow and collected in a can. Then several such cans are brought together in a creel area, and subtows from these cans are combined to feed a staple drawline. The latter may involve one or more stages of drawing and relaxation, one or more stuffer-box type crimpers, a drying unit, and a cutter. A baling unit following the cutter collects and bales the cut fiber.

Continuous Filament Yarn Process Variants

For producing continuous filament yarn, several process routes are available,[9] each of which involves the basic processes of polymer synthesis, extrusion, quenching, stretching, and winding. In one process, specially meant for textle type uses, spinnerettes with 20 to 100 holes typically are used. After solidification and finish application, the filaments may be split into smaller bundles, depending upon end-use applications, for downstream processing. The drawing can also be done in a single integrated process (spin-draw) immediately after finish application to yield a fully oriented yarn (FOY), which having no bulk or texture, is referred to as a flat yarn. Spinning speeds of the order of 1000 meters/min or more and winding speeds of about 4000 meters/min are used. Similar yarn may be produced on a separate drawing unit, but this process—called draw-twisting because a certain amount of twist is also inserted prior to winding—is less commonly used today. It is more usual to see the separate drawing process done in conjunction with a texturing process—most commonly false twisting, and called draw-texturing—to give a textured yarn. The process leads to orientation and crystallization of the filament structure and bulking of the yarn. The feed yarn for such processes usually is produced at 3000 to 4000 meters/min spinning speed as partially oriented yarn (POY). The latter is designed to be run on draw-texturing machines at speeds ranging from 800 to 1200 meters/min.

For industrial applications, polyester filaments having high strength low shrinkage (HSLS), low creep properties, and high glass transition temperatures are targeted. To produce such a filament, more severe processing conditions and higher molecular weight polymer are generally used. The filament is spun at low speeds (500–1000 m/min), sometimes with retarded quench, to obtain minimum orientation. The drawing can be achieved in the more common integrated, or spin-draw, process or in a separate draw-twist operation. High strength is achieved by drawing the filament to several times its length over very high-temperature rolls and then heat-setting and relaxing the structure prior to winding. Low shrinkage properties are obtained with a relaxing step at high temperature.

Modifications and New Developments

As was the case with the nylon fibers, the potential the polyester fibers offered in apparel, home furnishing, and industrial applications

was judged to be enormous. For this potential to be realized in practice, however, some characteristics had to be improved, and others had to be engineered for specific end uses. Thus, fibers of different cross-sectional shapes were developed in order to impart anti-soiling, reflective, and resilient characteristics for rug and carpet applications. A difficulty associated with the early polyester fiber that restricted its applications was its lack of ability to take on dyes through one of many methods available for dyeing. This problem was overcome by introducing chemicals that added sulfonate groups to the molecule and by substituting in some cases isophthalic acid for a small portion of terephthalic acid. These changes allowed fibers to be dyed by cationic and disperse dyes, the dyes most frequently used for polyester. Another area of modification has been the development of inherently flame-resistant fiber. One process involves copolymerizing a derivative of phosphoric acid with polyethylene terephthalate. An exciting new development in polyester filament yarn for apparel uses is the production of microdenier fiber (denier per filament less than 1), discussed later in a separate section. The introduction of finer-denier yarns opened up a whole new field for developing fabrics with special aesthetic and performance characteristics that were not possible earlier.

With the use of fibers in conveyor belts, tires, and composites, fibers of greater strength and modulus and lower extensibility have been needed. Much effort was directed in the 1970s, and later, to developing such fibers from polyester. The composition and the properties of wholly aromatic polyamides or aramids are discussed in a later section. When both the diacid and diamine components are *para*-substituted aromatic compounds, the resulting polymer is capable of forming lyotropic liquid–crystalline solutions. These solutions can be dry- or wet-spun into fibers with unusually high tensile strength and tensile modulus. When a similar strategy is tried to make polyester fiber from a homopolymer of a *para*-substituted aromatic diacid and a *para*-substituted aromatic diol, only infusible and intractable materials are obtained. A solution to this problem has been found in the development of polyester copolymers that give thermotropic liquid–crystalline melts over a useful temperature range and have viscosities suitable for melt extrusion into fibers or films having high levels of orientation. Spin-line stretch factors on the order of several hundred percent are used to achieve orientation, and physical properties are developed further by heat treatment at temperatures approaching melting conditions.

The first fibers from a thermotropic liquid crystalline melt whose properties were reported were spun from a copolyester of *para*-hydroxybenzoic acid (PHB) and poly-ethyleneterephthalate (PET) by workers at Tennessee Eastman Co. The preparation of the copolymer proceeds in two stages. First, *para*-acetoxybenzoic acid is reacted with poly-ethyleneterephthalate in an acidolysis step to give a copolyester prepolymer, which in the second step is condensed further to a higher degree of polymerization suitable for fiber formation.

When the mol% of PHB in the copolymer exceeds about 30 to 40 percent, a liquid–crystalline melt is obtained. Up to about 60 mol%, order in the melt increases and melt viscosity decreases. Compositions containing about 60 mol% PHB can be melt-spun into fibers using standard extrusion techniques. It is the unusual combination of properties that makes this class of materials valuable for the formation of high-strength fibers and plastics.

Among melt-spun fibers, those based on thermotropic liquid–crystalline melts have the highest strength and rigidity reported to date, and appear comparable to polyamides spun from lyotropic liquid–crystalline solutions. This was a very active field of research in the 1970s and later, and many comonomers have been reported. Obviously, these compositions must contain three components at a minimum, but many have four or five components. Some frequently used constituents, in addition to those mentioned above, are 2,6-naphthalene-dicarboxylic acid, hydroquinone, 4,4′-biphenol, isophthalic acid, and 4,4′-dihydroxy-diphenyl ether.

ACRYLICS

Polymer Manufacture

Acrylic fibers are spun from polymers that are made from monomers containing a minimum of 85 percent acrylonitrile. This compound may be made from hydrogen cyanide and ethylene oxide through the intermediate ethylene cyanohydrin:

$$\underset{\diagup\ \diagdown}{CH_2CH_2O} + HCN \longrightarrow HOCH_2CH_2CN \xrightarrow[-H_2O]{\text{catalyst}} CH_2{=}CHCN$$

It also may be made directly from acetylene and hydrogen cyanide:

$$CH{\equiv}CH + HCN \rightarrow CH_2{=}CHCN$$

But the reaction that currently is preferred uses propylene, ammonia, and air:

$$3CH_2{=}CHCH_3 + 3NH_3 + 7O_2$$

$$(\text{air}) \xrightarrow[<500°C]{\text{catalyst}} CH_2{=}CHCN + 2CO$$

$$+ CO_2 + CH_3CN + HCN + 10H_2O$$

Pure acrylonitrile may polymerize at room temperature to polyacrylonitrile (PAN), a compound that, unlike polyamides and polyesters, does not melt at elevated temperatures but only softens and finally discolors and decomposes. Nor is it soluble in inexpensive low-boiling organic solvents. Because fibers made from it resist the dyeing operations commonly used in the textile industry, the usual practice is to modify it by copolymerization with other monomers, for example, vinyl acetate, styrene, acrylic esters, acrylamide, or vinyl pyridine in amounts up to 15 percent of the total weight (beyond which the final product may not be termed an acrylic fiber). The choice of modifier depends on the characteristics that a given manufacturer considers important in a fiber, the availability and cost of the raw materials in the manufacturer's particular area of production, and the patent situation.

In copolymerizing acrylonitrile with another monomer, conditions must be controlled in such a way that the reaction produces a polymer having the desired chain structure and length. The reaction takes place in the presence of substances capable of producing free radicals. In addition, certain trace metals that have been found to increase reaction rates offer a means of controlling chain length. When polymerization is carried out in solution, after an induction period, the reaction is rapid and liberates a considerable amount of heat. Furthermore, because the polymer is not soluble in the monomer, a thick paste is formed. These facts limit the usefulness of such a process. Carrying out the polymerization in the presence of a large amount of water (water/monomer of 2/1 to 3/1) is a convenient method and the one most generally used. In this case the polymer forms a slurry, and the

water provides a means for removing the heat from the site of the reaction. Moreover, most of the common redox-catalyst systems are water-soluble. Polymerization may be carried out batchwise or by a continuous process.

In the batch method, the monomers and catalyst solutions are fed slowly into an agitated vessel containing a quantity of water. The heat of reaction is removed either by circulating cold water through the jacket surrounding the vessel or by operating the reaction mixture at reflux temperature and eliminating the heat through the condenser water. The monomer and catalyst feeds are stopped when the desired amounts have been added, and polymerization is allowed to continue until there is only a small amount of monomer remaining in the reaction mixture. Then the slurry is dumped from the reaction vessel, filtered, washed, and dried. (See Fig. 21.19.)

In the continuous-overflow method, rather than stopping the monomer and catalyst feed when the reaction vessel is full, the slurry is simply allowed to overflow; the solids are removed by filtration, washed, and dried. The filtrate contains a certain amount of unreacted monomer, which is recovered by steam distillation after the trace metal present has been chelated to stop the redox reaction and thus further polymerization. The dried polymer is the raw material from which fibers are spun.

As mentioned, polyacrylonitrile polymer has two major weaknesses: (1) extremely poor dyeing characteristics due to highly ordered structure, and (2) very low solubility in most of the solvents. To overcome these problems, comonomers are incorporated into the chains. Comonomers have been used to improve hydrophilicity, dyeability, flame retardancy, heat resistance, and so on. For obtaining hydrophilic acrylic fiber, comonomers containing hydrophilic functional groups such as hydroxyl, ester, carboxyl, amide, and substituted amide are used. To make the fiber base-dyeable, comonomer containing carboxylic and sulfonic acid groups is introduced; and to obtain acid-dyeable fiber, comonomers containing pyridine, aliphatic amine, or quaternary ammonium salt are used. Antistatic acrylic fibers can be made by incorporating in the chain polyethylene oxide, polyalkylene derivatives, polyethylene glycol, acrylates, or methacrylates as the comonomers. For improved flame retardancy, halogen-containing vinyl comonomers are used.

Spinning

As already indicated, pure polyacrylonitrile softens at elevated temperatures, and thermal decomposition starts before the molten state is reached. The same is true of the copolymers

Fig. 21.19. Flow diagram for the manufacture of acrylic fiber: (1) acrylonitrile; (2) tank farm; (3) polymerizer; (4) comonomer and catalyst; (5) centrifuge; (6) waste liquid; (7) dried polymer; (8) grinding; (9) polymer storage; (10) dissolver; (11) filter; (12) solvent plant; (13) spinnerette; (13w) wet spinning; (13d) dry spinning; (14) roller dryer; (15) additional treatment; (16) crimper; (17) cutter; (18) acrylic fiber bale.

commonly used to produce fibers. Accordingly, melt spinning is impossible; spinning must be done from a solution of the polymer. Both dry and wet spinning are carried out in current commercial operations.

The operations used to either wet- or dry-spin acrylics are essentially the same as those already described for rayon and acetate, respectively. The polymer must be completely dissolved in solvent and the solution filtered to remove any impurities that would cause spinnerette blockage. Because acrylic polymers are not soluble in common nonpolar solvents, polar substances such as dimethylformamide, dimethylacetamide, or aqueous solutions of inorganic salts such as zinc chloride or sodium thiocyanate are required. Only wet spinning is possible with the latter. Dimethyl formamide boils at 152.8°C and exerts a vapor pressure of 3.7 mm of Hg at 25°C compared with acetone (used in dry spinning of cellulose acetate), which has a vapor pressure of 228.2 mm of Hg at 25°C. It follows that, unlike acetone which requires an activated-carbon system for recovery, dimethylformamide may be condensed directly from the gas stream used to evaporate the solvent from the forming fiber.

In order to obtain the desired characteristics of modulus, rupture tenacity, and rupture elongation, acrylics, like rayon, require stretching which is usually carried out after the fiber has been spun, either as part of the fiber washing operation or after the fiber has been dried. These same properties are influenced by spinning speeds, and the temperature of the drying air, if they are dry-spun, or the temperature and the composition of the bath, if wet-spun. The multitude of combinations made possible by the use of various comonomers and the flexibility of the fiber-forming operations furnish the different manufacturers with versatility and the users with a variety of acrylic fibers. Figure 21.20 shows a wet-spinning operation for acrylic tow.

Acrylic fibers possess a property that made it possible for them in the late 1950s and early 1960s to find immediate, even spectacular, acceptance in the knitted sweater field, until then dominated by wool. When acrylic fibers, normally in the form of a heavy tow, are hot-stretched (e.g., by being drawn over a hot

Fig. 21.20. Wet spinning of acrylic tow. (*Courtesy Monsanto Co.*)

plate and then cooled under tension), they are converted to a labile state. Upon immersion in hot water such fibers will contract considerably, but not to their prior unstretched length. In practice this characteristic is used to produce a bulky yarn resembling the woolen yarns long accepted for use in sweaters. The process is described briefly below.

Using "stretch-break" equipment, the stretched labile fibers are further cold-stretched to the breaking point so that the fiber breaks at different points leading to a distribution of fiber lengths, similar to the lengths found in wool. These are crimped and then mixed with thermally stable acrylic fibers that have been stretched and relaxed and have about the same length and degree of crimp. The blend is converted to a spun yarn by the same process used in making woolen yarns, and in turn this yarn is knitted into sweaters and other similar products. When such garments are dyed in hot water, the labile fibers, intimately blended with stable ones, contract lengthwise individually. In the process segments of the stable units tend to be carried along physically by entrapment and friction; but because such fibers do not change their overall length, the yarn as a whole decreases in length. Lateral displacement of the large volume of stable fibers results in the formation of a more voluminous structure known as a "bulky" or "hi-bulk" yarn.

Bicomponent or Conjugate Spun Fibers

As will be shown, it should be theoretically possible to make any of the common manufactured fibers in bicomponent forms. However, acrylics have received the most attention for quite good reasons. Their general characteristics have tended to make them competitive with wool. This means that they should be processible on machinery developed for handling wool, as well as capable of being accepted into markets previously dominated by wool. It follows that because the natural fiber possesses crimp which produces the cohesion that determines its behavior in processing and in part its appearance and "hand" in usage, a similar crimp was desired for acrylics.

The principle that is the basis for bicomponent fibers usually is likened to that which underlies the bicomponent metal strips often used in temperature controllers. With the latter, differential-thermal expansion of the two joined components results in a bending of the thermal element. With fibers, moisture usually is the agent that acts upon the two side-by-side portions. Differential swelling or shrinkage causes the fiber to be brought into a crimped, or preferably, a spirally distorted condition. Incidentally, the side-by-side structure exists naturally in wool.

The combination of small size and large number of holes in a spinnerette might lead one to conclude that it would be almost impossible to design a spinnerette assembly that could bring two streams of polymer or polymer solutions together at each such hole and extrude them side by side to form a single filament. Such designs have, in fact, been made; but solutions of fiber-forming polymers fortunately possess properties that encourage laminar flow and thus make other approaches possible. This phenomenon was remarked upon earlier in connection with dope dyeing; when a suspension of a colored pigment is injected into a dope stream, a considerable problem must be overcome to achieve adequate mixing so as to secure "dope dyed" fibers of a uniform color. Thus, it was known that when two streams of essentially the same solution of a fiber-forming polymer are brought together, side-by-side, and moved forward down a pipe or channel by the same amount of pressure behind each, virtually no mixing takes place. By bringing these streams to each spinnerette hole in such an individual side-by-side arrangement and using appropriate mechanical separators, the extruded filament from each hole will have a bicomponent structure. In addition to producing fibers in which the two components form a bilateral symmetrical structure, an ingenious arrangement of predividers of the two streams can produce from the full complement of holes in a single spinnerette a selected group of fibers wherein the amount and the position of each

of the two components are randomly distributed throughout their cross sections. It follows that curls of uniform or random geometry may be produced to meet required needs.

VINYL AND MODACRYLIC FIBERS

Vinyls

When nylon 66 was developed, it was described as being "synthetic" or "fully synthetic" in order to differentiate it from rayon and acetate. This was no small act of courage, as the word "synthetic," in that period just following the repeal of Prohibition in the United States, was often associated in the public mind with the least palatable kind of alcoholic beverages. In due time, what is known in the advertising business as "puffing" led it to be known as the "first fully synthetic fiber," which was an anachronism. It so happens that fibers based upon polyvinyl chloride predated nylon by several years.

About 1931, the production of fibers from polyvinyl chloride (PVC) was accomplished by dry spinning from a solution in cyclohexanone. But by chlorinating the polymer, it was possible to secure solubility in acetone, which has the advantage of possessing a boiling point about 100°C lower than that of cyclohexanone. Several million pounds per year of this fiber were produced in Germany during World War II to relieve the shortages of other materials. Unfortunately, PVC begins to soften at about 65°C, and in the fibrous state it shrinks disastrously upon heating. Because of its low softening point, it cannot be dyed at the temperatures commonly used for this purpose, and, furthermore, it resists dyeing.

Modifications of PVC have been produced by copolymerization with other monomers. The first successful one consisted of 90 percent vinyl chloride copolymerized with 10 percent vinyl acetate. It was dry-spun from acetone and given the trade name Vinyon by its producer, Union Carbide Corporation. (In 1960 vinyon was accepted as a generic name for fibers containing not less than 85 percent vinyl chloride.) It has never been produced in large volume; it is used for heat-sealable compositions.

A copolymer of vinyl chloride with vinylidene chloride was used for a number of years to produce melt-spun, heavy monofilaments, which found use in heavy fabrics where the chemical inertness of the polymer was needed, in outdoor furniture, and in upholstery for seats in public-transportation vehicles.

Another vinyl-based fiber, polyvinyl alcohol, or vinal, was developed in Japan but has not been produced or used in the United States. As such, it illustrates the importance of both relative availability of raw materials and differences in markets, in the success of a chemical product. Acetylene made from calcium carbide is converted to vinyl acetate, which, following polymerization, is saponified to polyvinyl alcohol.

The polyvinyl alcohol is soluble in hot water, and the solution is wet-spun into a coagulating bath consisting of a concentrated solution of sodium sulfate. The fibers are heat-treated to provide temporary stability so that they may be converted to the formal derivative by treatment with an aqueous solution of formaldehyde and sulfuric acid. This final product resists hydrolysis up to the boiling point of water. It seems reasonable to assume that it contains hemiacetal groups and some unreacted hydroxyls on the polymer chain as well as cross-linking acetyl groups between the adjacent molecules.

Under the trade name Kuralon® (Kuraray Co., Ltd), it achieved a production level of about 180 million lb in 1970, but production

dropped to 16 million lb of continuous filament and 87 million lb of staple in 1980. The former has been mainly used in industrial rubber products, and the latter has been used mostly for uniforms, nonwoven and coated fabrics, and filters.

Modacrylics

In the United States, the modification of PCV has moved in the direction of copolymerizing vinyl chloride with acrylonitrile, or perhaps it should be said that PAN has been modified by copolymerizing the acrylonitrile with chlorine-containing vinyl compounds. In any case, one modacrylic fiber is currently produced in the United States, a modacrylic being defined as containing at least 35 percent but not over 85 percent acrylonitrile.

The first two modacrylic fibers ever introduced in the United States were Dynel® (by Union Carbide) in 1949 and Verel® (by Tennessee Eastman) in 1956. The former was a copolymer of 60 percent vinyl chloride and 40 percent acrylonitrile, and the latter was said to be a 50–50 copolymer of vinylidene chloride and acrylonitrile with perhaps a third component graft-co-polymerized onto the primary material to secure dyeability. SEF® and its version for wigs, Elura®, were introduced by Monsanto Fibers in 1972. A few foreign manufacturers are making modacrylic fibers, but the only modacrylic fiber currently in production in the United States is SEF®.

Modacrylic fibers, like acrylic, require after-stretching and heat stabilization in order to develop the necessary properties. It is thought that the stretching is on the order of 900 to 1300 percent, and that, in a separate operation, shrinkage of about 15 to 25 percent is allowed during the time that the fibers are heat-stabilized.

The modacrylic fibers, like vinyon and unlike the acrylic fibers, have not become general purpose fibers. They can be dyed satisfactorily and thus are acceptable in many normal textile products; but their nonflammability tends to place them in uses where that property is important, even vital. Blended with other fibers, they are used in carpets; but their largest market is in deep-pile products, such as "fake-furs," or in doll hair, where a fire hazard cannot be tolerated.

ELASTOMERIC FIBERS

The well-known elastic properties of natural rubber early led to processes for preparing it in forms that could be incorporated into fabrics for garments. One such process uses standard rubber technology. A raw rubber of high quality is compounded with sulfur and other necessary chemicals, calendered as a uniform thin sheet onto a large metal drum, and vulcanized under water. The resulting skin is spirally cut into strips that may be as narrow as they are thick, for example, 0.010 inch by 0.010 inch square in cross section. These strips are desulfurized, washed, dried, and packaged. Larger cross sections are easier to make. This product, coming out of the rubber rather than the textile industry, is known as a thread.

Another method produces a monofilament known as a latex thread. As the name would indicate, rubber latex is the raw material, and because extrusion through small holes is required, the purity of the material must be of a high order. With proper stabilization, the latex solution may be shipped from the rubber plantation to the plant, where it is compounded with sulfur and other chemicals needed for curing, as well as with pigments, antioxidants, and similar additives. This is followed by "precuring" to convert the latex to a form that will coagulate upon extrusion into a precipitating bath of dilute acetic acid and will form a filament having sufficient strength for subsequent operations. It passes out of the bath and is washed, dried, vulcanized in one or two stages, and packaged.

The rubber threads manufactured by either process can be used as such in combination with normal nonelastomeric yarns in fabrics made by weaving or knitting; but most of them, especially those made by a latex process, first are covered by a spiral winding of natural or manufactured yarns. Often two layers are

applied in opposite directions to minimize the effects of torque. Such coverings have two purposes. The first is to replace the less desirable "feel" of rubber on human skin by that of the more acceptable "hard" fiber. The second concerns the engineering of desired properties into the product to be woven or knitted into fabric. As an elastomeric material begins to recover from a state of high elongation, it supplies a high stress; but as it approaches its original unstretched condition, the stress drops to a very low order. When wound in an elongated state with a yarn having high initial modulus and strength, the elastomeric component cannot retract completely because its lateral expansion is limited, and jamming of the winding yarn occurs. Thus the combination of such materials can be made to provide stretch and recovery characteristics needed for a broad spectrum of applications.

The traditional elastomeric threads have been subject to certain inherent limitations, however. The presence of unreacted double bonds makes them sensitive to oxidation, especially with exposure to the ultraviolet radiation of direct sunlight. They also have low resistance to laundry and household bleaches and dry-cleaing fluids.

During recent years, elastomeric yarns or threads have been used to impart comfort, fit, and shape retention to a variety of garments such as women's hosiery and swimwear. Such garments must be thin and highly effective per unit of weight. The materials of which they are composed must be compatible with these requirements. Thus, it was not unexpected that the producers of manufactured fibers, already eminently successful in meeting the needs of the marketplace, should look to the field of elastomeric fibers for new possibilities. Given the limitations of rubbers, both natural and synthetic, as well as the relationships between molecular structure and behavior of fiber-forming linear polymers, the scientists faced new challenges.

As an oversimplification, it can be said that within limits a rubber-like material can be stretched relatively easily but reaches a state where crystallization tends to occur. The structure produced in this manner resists further extension, and the modulus rises sharply. In contrast to the conditions that occur when the manufactured fibers discussed earlier such as nylon or polyester are drawn to form fibers of stable geometry in the crystalline and oriented states, the crystalline state of the elastomeric fibers is labile unless the temperature is lowered materially. Thus, to improve on the chemical sensitivity of rubber, new approaches were necessary. The solution was found in developing linear block copolymers containing "soft" liquid-like sections that impart elasticity, connected with "hard" components that act as tie points to hold the structure together.

The soft, flexible, and low-melting part is commonly an aliphatic polyether or a polyester with hydroxyl end groups and molecular weight in the range of 500 to 4000. The hard portion is derived from an aromatic diisocyanate supplied in an amount that will react with both end groups of the polyether or polyester to form urethane groups. The product, an intermediate known as a prepolymer, is a thick liquid composed essentially of molecules carrying active isocyanate groups at each end. For example:

$$\mathrm{HO}\text{-}(\mathrm{RO})_n\text{-}\mathrm{H} + m\,\mathrm{OCNR'NCO} \rightarrow \mathrm{OCNR'N}\overset{\overset{\mathrm{HO}}{\|}}{\mathrm{C}}\mathrm{O}\text{-}(\mathrm{RO})_n\text{-}\overset{\overset{\mathrm{OH}}{|}}{\mathrm{C}}\mathrm{NR'NCO}$$

where $\text{-}(\mathrm{RO})\text{-}$ is an aliphatic polyether chain, R′ is one of several commonly available ring structures, $n \sim 10$–30, and $m \sim 1.5$–2.

The elastomeric polymer is obtained by "extending" the prepolymer through its reaction with short-chain diols such as butanediol or diamines such as ethylene diamine, thus completing the formation of hard groups between soft, flexible chains. When amines are used, the final step is typically done in a polar solvent such as dimethyl acetamide. The conversion of these polymers into usable fibers may be accomplished by wet-, dry-, or melt-spinning operations, depending on the polymer. Additives to impart whiteness or

improve resistance to ultraviolet radiation and oxidation may be incorporated in the spinning solutions or in the melts.

The development of elastomeric fibers has resulted in a variant of wet spinning called "reaction" or "chemical" spinning. In point of fact, rayon, the first wet-spun material, might properly be said to be produced by "reaction wet spinning" or "chemical wet spinning" because complex chemical reactions always have been involved in that operation. In any case, it has been found that the prepolymer of an elastomeric fiber may be extruded into a bath containing a highly reactive diamine so that the chemical conversion from liquid to solid occurs there.

The elastomeric fibers produced in this fashion are based upon segmented polyurethanes and by definition are known generically as spandex yarns. Each manufacturer uses a trade name, for the usual commercial reasons. Perhaps the most noteworthy aspect from the standpoint of industrial chemistry is the multitude of options available to the manufacturer through the ingenious use of various chemicals for soft segments, hard units, chain extenders, and conditions of chemical reaction, followed by numerous possibilities for extrusion and after-treatments. In the United States, there are two main producers of spandex fibers: Du Pont (Lycra®) and Globe Rubber Co. (Cleerspan®, Glospan®). There are numerous worldwide producers, including: Bayer, Germany (Dorlastan®); Asahi, Japan (Roica®); Nisshinbo, Japan (Mobilon®); and Tae Kwang, Korea (Acelan®).

POLYOLEFIN FIBERS

Polypropylene

Although polyethylene was considered a source of useful fibers at an early date, its low melting point (110–120°C) as well as other limitations precluded active development during the period when production of other fibers based upon the petrochemical industry expanded enormously. The higher melting point of high-density polyethylene gave some promise, but it was overshadowed by the introduction of polypropylene (PP) around 1958–59. Great expectations were held for the latter as a quick competitor with the polyamides and the polyesters, already successful, as well as the acrylics, which then were entering the fiber field in volume. Polypropylene was thought to have several advantages. The raw material costs were low, only a few cents a pound; also there was a high level of sophistication in the spinning and processing of fibers, and a presumption that this would readily lead to the development of means for converting the polymer to fibers; and, finally, there was the belief that the American consumer would be ready to accept, and perhaps even demand, something new and different, which this polymer offered. However, the limitations of polypropylene fibers, such as lack of dyeability, low melting temperature, low heat stability, and poor light stability, combined with the lower prices and the greater versatility of the already established fibers, dashed the hopes for quick success. However, all of these deficiencies except the low melting temperature and lack of dyeability now have been overcome. The fiber has found an increasingly important place, and its properties have led to new techniques of manufacture and specialized uses.

The structural formula of PP is as given below, where $100{,}000 < n < 600{,}000$ for chips or granules, and $50{,}000 < n < 250{,}000$ for fibers:

$$\left[-CH_2-\underset{\displaystyle CH_3}{\underset{|}{CH}}- \right]_n$$

The steric configuration is extremely important in the polymer. Only isotactic polypropylene (iPP) has the properties necessary for forming fibers. The molecules are cross-linked only by Van der Waals forces, so it is important that they pack as closely as possible. The isotactic molecules form a 3_1 helix, as shown in Fig. 21.21,[10] and exhibit a high crystallization rate. The atactic molecules, shown in the figure, do not pack well, and although the syndiotactic molecules can pack better and crystallize, this

Fig. 21.21. Diagrams showing irregular atactic (a), stereoregular isotactic (b), and stereoregular syndiotactic (c) configurations in polypropylene polymer. (Source: *Ahmed, M., "Polypropylene Fibers—Science and Technology,"* Textile Science and Technology 5, *p. 16, Elsevier Science Publishers SV, Academic Publishing Div., New York, Amsterdam, 1982.*)

configuration is not a normal product of commonly used catalyst systems. Some properties of isotactic, syndiotactic, and atactic PP are listed in Table 21.1.[11] The insolubility of isotactic polypropylene in hydrocarbon solvents at room temperature often is used to separate iPP from atactic polypropylene (aPP).

Early in the manufacture of polypropylene a concept was developed for dry spinning directly from the solution obtained in the polymerization operation. Had it been feasible, it would have been the realization of a chemical engineer's dream: the gaseous olefin fed into one end of the equipment, and the packaged fiber, ready for shipment to a textile mill, coming out the other end. But it did not turn out that way, and today melt spinning is the accepted technique for the production of staple fibers, monofilament, and multifilament yarns. To this usual method have been added the fibrillation and the "slit film" procedures for producing yarns.

The polypropylene materials are completely resistant to bacterial attack, are chemically inert, and are unaffected by water. Monofilaments can be produced that possess high strength, low elongation under stress, and dimensional stability at normal atmospheric temperatures. Polypropylene monofilaments have found broad application in cordage and fishing nets (which float), and if highly stabilized they are woven into fabrics used for outdoor furniture, tarpaulins, and similar applications. Large filament denier staple is used widely in "indoor–outdoor" carpets. Also, staple fibers have found major applications in tufted indoor carpets and nonwovens used for diaper, filtration, and civil engineering fabrics.

Synthesis. The early polypropylene plants used a slurry process adopted from polyethylene technology. An inert liquid hydrocarbon diluent, such as hexane, was stirred in an autoclave at temperatures and pressures sufficient to keep 10 to 20 percent of the propylene monomer concentrated in the liquid phase. The traditional catalyst system was the crystalline, violet form of $TiCl_3$ and $AlCl(C_2H_5)_2$. Isotactic polymer particles

TABLE 21.1 Properties of Isotactic, Syndiotactic, and Atactic Polypropylene

Property	*Isotactic*	*Syndiotactic*	*Atactic*
density, g/cm^3	0.92–0.94	0.89–0.91	0.85–0.90
melting point, °C	165	135	
solubility in hydrocarbons at 20°C	none	medium	high
yield strength	high	medium	very low

Source: Lieberman, R. B., and Barbe, P. C., "Propylene Polymers," in *Concise Encyclopedia of Polymer Science and Engineering*, p. 916, J. I. Kroschwitz (Ed.) (1990). Copyright © John Wiley and Sons and reproduced by permission of the copyright owner.

that were formed remained in suspension and were removed as a 20 to 40 percent solid slurry while the atactic portion remained as a solution in the liquid hydrocarbon. The catalyst was deactivated and solubilized by adding HCl and alcohol. The iPP was removed by centrifuging, filtration, or aqueous extraction, and the atactic portion was recovered by evaporation of the solvent. The first plants were inefficient because of low catalyst productivity and low crystalline yields. With some modifications to the catalyst system, basically the same process is in use today.

In 1963, liquid polymerization was introduced in which liquid propylene, catalysts, and hydrogen were pumped continuously into the reactor while polypropylene slurry was transferred to a cyclone separator. The unconverted monomer gas was removed, compressed, condensed, and recycled, and the polymer was treated to reduce the catalyst residue. This system also suffered from a poor catalyst yield, and the polymer produced lacked the required stereospecificity, so that it was necessary to remove the atactic portion of the polymer.

In the mid-1960s a gas phase process was introduced for production of the polymer. The monomer was pumped over adsorbing beds and entered the reactor with the catalyst system. These feed streams of monomer and catalyst, together with a mechanical stirrer, created a turbulent bed of powdered polymer. Periodically the polymer powder was vented off in a carrier gas to extrusion storage hoppers. Meanwhile, the heat of polymerization was removed by condensing the unreacted monomer in a cooling loop and returning it to the reactor, where it immediately vaporized. This process eventually led to the production of highly crystalline products and was adopted by several companies in the United States.

Most processes in use today rely on a combination of these technologies. Montedison's introduction in 1975 of third-generation catalysts gave high yields and allowed polymerization to take place at 60 to 80°C and 2.5 to 3.5 MPa (362–507 psi). This was welcome news during the energy crisis, but the resulting polymer was not stereospecific enough to eliminate the need for removal of aPP. Real progress came with the discovery of superactive third-generation catalysts, which gave both the optimal yield and stereospecificity.[12]

Production. Classical melt spinning, which was developed for the production of nylon filaments, is widely used to produce polypropylene fibers today. It involves a high-speed process (2000–3000 meters/min) that is particularly suitable for long production runs. The average molecular weight of polypropylene polymer, like that of other addition (olefin) polymers, is relatively high compared to that of other polymers. This results in a high melt viscosity; so, unlike the case of other polymers, its extrusion temperatures are 70 to 100°C above its melting point. Single-screw extruders are used for melting and homogenizing the polymer. The screw diameters are from 45 to 200 mm and screw lengths are 24, 30, or 36 times the diameter. The polymer granules are fed into the extruder hopper, where they are melted and homogenized. Chips carrying pigments can be fed into and blended with the main charge of the extruder if colored fibers are desired. The molten polymer is forced through the spinnerette via a screen pack to eliminate any contaminant particles. The spinnerette hole determines the shape of the filaments, and the flow rate and the takeup speed determine the size. The polymer has high specific heat and low thermal conductivity, so the cooling zones must be longer than those for polyester and nylon. For filament yarn production, the filaments are drawn at high speed and wound on packages. For staple fiber production, the filaments are collected in the form of tow and then are drawn, crimped, and cut. The multifilament yarns are often textured to improve bulk and appearance. The false twist method is generally used to texture finer yarns, with the stuffer box used for coarser yarns.

The short spinning method used to produce staple fiber is considerably slower than high-speed spinning. The lower spinning speeds (30–150 meters/min) would have a negative

effect on productivity, but this is counteracted in industry by the use of spinnerettes with a large number of holes (up to 55,000). The required cooling zones are much shorter because of the lower speeds and the use of higher volumes of quench air, which gives this method its name. Because the drawing units can match these low speeds, the two can be fed directly and continuously from the spinning machine to the draw frame to the texturing chamber or the crimper, to produce bulked continuous filament (BCF) yarn or staple fiber, respectively. The short spinning method is used to produce high-tenacity fibers. Many polypropylene yarns are produced using the slit film method. The film extruder is almost identical to a filament extruder. In it, the molten polymer is forced through a film die that converts the melt into film, where the thickness of the film can be controlled by adjusting one of the die lips. The takeoff unit is either a chill roll that removes the film uniformly and cools it below T_m or a water bath followed by nip rolls. The cooled film is slit into separate tapes using a slitter bar that contains a large number of special knives separated by spacers. The film tapes are heated and drawn to their final length.

Several methods are used to produce fibrillated film. They are produced in much the same way as slit film, but these techniques take advantage of the tendency of PP to fibrillate. In one method a profiled tape is extruded and hot-drawn. It is drawn again to achieve a 10:1 ratio, and the film splits into separate filaments. In another method, called roll embossing, the film is hot-drawn, and then embossed using profiled rollers. The profiled film is drawn again, and fibrillation results. In the pin-roller technique, the film is drawn and cut by knives or pins on a rotating cylinder. This method can produce individual fibers or a controlled web network.

Polypropylene nonwovens are created by forming a staple fiberweb and then consolidating it into a fabric. The fibers can be entangled by a needling machine, an air jet, or a water jet. Another method uses a single-stage process in which melt-spun fibers are drawn through an air aspirator jet and deposited randomly on a conveyor. The fibers then are bonded by fusion under heat and pressure to give a "spunbonded" nonwoven fabric (see Fig. 21.22). Yet another method, known as melt blowing, can be used in producing a nonwoven web having special characteristics. The melt-blown fibers are characterized as ultrafine fibers because of their size relative to other fibers. In the process, molten polymer is forced through a melt blowing die and die tip orifice that are finer than most orifices used to make manufactured fibers. As the polymer emerges, it is attenuated by a jet of high-velocity hot air. This allows the polymer to stay in a molten state but attenuate until broken. The fibers then come in contact with cool quench air, which causes the former to solidify. The fibers deposit on a collector screen and form a

Fig. 21.22. Spunbonded polypropylene showing interfiber bonding, which binds the structure. (*Courtesy E. I. du Pont de Nemours and Co.*).

melt-brown web. These webs, composed of very fine size fibers, can be engineered for applications in areas where barrier (such as against bacteria), insulative, and absorptive (such as for oil) characteristics are important.

The production figures of polypropylene show impressive growth. The volume, which was less than 10 million lb in 1965, reached over 1.5 billion lb in 1990. This included approximately 17 percent filament yarn, 23 percent staple fiber, 18 percent spunbonded fabrics, and 38 percent split film products.

High Performance Polyethylene

A recent addition to the group of high performance fibers is the ultra-high-molecular-weight, extended-chain, linear polyethylene. Although fundamental work in the area of developing fully oriented and crystalline structures in polyethylene polymer had been going on since the mid-1960s, it was not until the late 1970s that the possibility of producing such materials on a commercial scale became evident. Presently, three companies worldwide are manufacturing extended-chain polyethylene fibers, which have very high strength and high modulus characteristics.

Routes to High Performance. The achievement of ultimate strength and high modulus in fibers has been a subject of great interest to material scientists and fiber producers. In an attempt to identify the preferred structures for high performance, many theoretical analyses have been conducted to calculate the limiting values. The basis of such calculations is the assumption that in a fiber in which all polymer chains of infinitely long dimensions are extended and oriented parallel to the axis, rupture will occur only when the stress exerted exceeds the intramolecular bond strength. Such calculations show that the limiting values of conventional fibers usually are several times higher than those obtained in actual practice. The reasons for this difference lie in one or more of the following explanations: (1) the molecular weights are not high enough; (2) the chains are not fully extended; (3) the chains are not fully oriented. For a given polymer system and molecular weight, tensile strength and modulus can be enhanced by extending and orienting the chains. Practically, many modern techniques, including zone stretching, multiple-step stretching, and state-of-the-art high-speed extrusion methods, have been adopted to achieve such results.

Still, the results of studies on conventional fibers show that modifications in physical processing alone cannot lead to values that even approach the theoretical maximum within an order of magnitude. Flexible chains, which characterize the bulk of the commercial polymers, tend to conform to a random coil or folded chain structure in an as-spun material, and are very difficult to reorganize into an extended-chain structure by known methods. Accordingly, in order to achieve ultimate properties, either novel spinning methods are needed which allow flexible chains to be fully extended, or the chains used must be so configured that they have high intrinsic stiffness and would remain extended in solution or a melt. Both of these goals have been met and are exemplified in the production of (1) high-strength polyethylene fibers, discussed in this section, and (2) aramid fibers, discussed in the next section.

Extended-chain polyethylene fiber became available commercially in 1984 when DSM, a Dutch firm in the Netherlands, introduced Dyneema®, and Mitsui Petrochemicals in Japan announced Tekmilon®. Allied Signal of the United States entered the field in 1985 when it introduced Spectra® fibers. These materials are characterized by very high strength and modulus, which are achieved by the use of ultra-high-molecular-weight polyethylene spun by the gel spinning method into fibers having extended-chain structures and near perfect orientation.

Gel Spinning. In general the purpose for which the gel spinning method is used is to produce an as-spun fiber that contains a loose network of chains with few entanglements, which then can be drawn out to ultrahigh levels to yield a highly oriented structure. The surface-growth method of Pennings,[13] which uses a Couette type apparatus, and the gel

spinning method of Smith-Lemstra,[14] which uses a more conventional spinning apparatus, led to the achievement of this goal. In the former method, a polymer solution is stirred between two counterrotating cylinders that provides the elongational flow necessary for initial chain alignment. Essentially the same result is achieved by passing the polymer solution through a constriction prior to spinning, in the latter method.

In the surface-growth technique, polymer solution between the inner rotating cylinder and the outer stationary cylinder is maintained within a certain temperature range above the polymer crystallization temperature. A fiber seed is immersed through an opening in the outer cylinder, and its tip is made to attach to the polymer layers absorbed on the inner rotating cylinder. By pulling on the other end of the fiber and winding it onto a bobbin, while at the same time replenishing the solution in the gap between the cylinders, a fiber can be produced continuously. This process results in a fibrous precipitate with a "shish-kebab" morphology (Fig. 21.23). One can conclude, then, that the flow field that is developed extends the chains, which then crystallize in a fibrillar form. Chains that are left unextended, those that are below a certain chain length for the speed, use the preformed clusters (shishes) as nuclei and crystallize as overgrowths in the form of chain-folded lamellae (kebabs). Fibers formed in this way showed remarkable mechanical properties with modulus in excess of 1200 gram force/denier (102 GPa) and tensile strength in the neighborhood of 30 gram force/denier (2.6 GPa).

Production. Recognition that the shish-kebab fibers produced by the surface-growth procedure result from the deformation of a gel-like entangled network layer at the rotor surface led to the development of gel-spun polyethylene fibers. The fiber is made by the solution spinning method. The polymer is dissolved in a high-temperature solvent, such as decalin, at a low concentration, 10 percent or less, and extruded into cold water. Initially the fiber is formed as a gel that still contains the spinning solvent. It can be oriented by drawing at remarkably high draw ratios of 30 or greater, either before or after solvent removal. A method of spinning PE fiber is described in the patent by Kavesh and Prevorsek.[15] The factors that govern the properties of fibers produced by this method are the polymer molecular weight, the concentration of polymer in the solvent, the type of solvent, the solution temperature, and the stage and the extent of drawing.

The high-performance behavior in these fibers is obtained from their having a very high molecular weight (1–5 million) combined with a very high degree of extended-chain crystal continuity. The fiber can aproach a 100 percent crystalline structure with a theoretical maximum density ($\sim$0.97 g/cc). Their use of course, is, limited by the melting point of the polymer, which even in the extended form is only about 150°C. Although higher than that of ordinary polyethylene, it is still much lower (80°C or more) than those of commercial textile and other higher performance industrial fibers. There also has been concern about the creep that occurs in these fibers, although significant improvement

Fig. 21.23. "Shish-kebab" structure showing extended-chain crystals with lamellar overgrowths.

has been made in this regard since the introduction of the first fiber. However, these fibers have such a unique combination of strength and lightness that they have proved highly successful in a number of applications, such as sailcloth, body armor, medical implants, fishing net, and sports equipment.

ARAMIDS

Introduction

As pointed out in the preceding section, a second route for developing fibers having properties approaching the ultimate is the use of polymer chains that have high intrinsic stiffness and will remain extended in solution or melt. The development of aramid organic fibers based on aromatic polyamides met these requirements and added another chapter to the history of the development of synthetic fibers. Nomex® aramid, a thermally resistant fiber based on a *meta*-oriented structure, was commercialized by the Du Pont company in 1962.

Following the technological breakthroughs which led to the discovery of (1) the liquid crystalline behavior of *para*-oriented aramids[16] and (2) a novel method for spinning anisotropic liquid crystalline polymer solutions,[17] Kevlar® aramid fiber was produced and commercialized by the Du Pont company in 1972. Other fibers based on aromatic polyamide compositions, which were produced and commercialized by other companies, were Technora® (Teijin, Japan), Teijinconex® (Teijin, Japan), and Twaron® (Akzo, The Netherlands). Additionally, SVM is a fiber produced in the Former Soviet Union and it was announced in 1990 that a new aramid fiber had been introduced by Hoechst, in Germany.

The rigid rod chains in *para*-aramids tend to form so-called liquid crystals when they are dissolved in polar solvents or heated to certain temperatures. The polymers showing liquid crystalline behavior in melts are called the thermotropic liquid crystalline polymers, and those showing similar properties in solution are called the lyotropic liquid crystalline polymers. These liquid crystals exist in three distinct phases according to their specific structures (Fig. 21.24).

A one-dimensional arrangement of rigid rods is called a nematic structure; a two-dimensional ordered arrangement represents a smectic structure; and the cholesteric structure is formed in terms of rotating oriented sheets so that rigid rods align parallel to each other in every layer, but the directional vectors in each layer are different. The preferential phase considered for fiber spinning is nematic. Nematic solutions or melts are easy to develop into oriented structures

Nematic Smectic Cholesteric

Fig. 21.24. Structure of liquid crystalline polymers showing nematic, smectic, and cholesteric orders.

through shearing and elongational flow during extrusion because of the rigidity of the polymer chains. Thus high orientation can be obtained in the as-spun fibers without much post-treatment.

Manufacture

In 1973, the Federal Trade Commission recognized aramid as a distinctly different generic material and defined it as "a long-chain synthetic polyamide in which at least 85% of the amide linkages are attached directly to two aromatic rings." This distinguishes aramids from nylon, which was redefined as a polyamide with less than 85 percent of the amide linkages attached to two aromatic rings. The first aramid fiber produced in the United States was Nomex®, the reactants being *m*-phenylene diamine and isophthaloyl chloride to give poly(*m*-phenylene-isophthalamide) (MPD-I).

MPD-1

The physical and chemical properties of this fiber are not remarkably different from those of other strong polyamides, but it does have excellent heat and flame resistance that makes it particularly suited for use in protective clothing and in specified industrial end uses. Military flight suits, fire-fighter uniforms, and hot gas filtration are a few of its many possible applications.

The other important fiber in this category, which also was first produced commercially in the United States, is Kevlar®, introduced in 1971 as fiber B and later coded as Kevlar®29. It was produced from poly(*p*-phenylene terephthalamide).

poly(*p*-phenylene terephthalamide)

Later, a higher-modulus version, Kevlar®-49, believed to be made by heat anealing of Kevlar®29, was introduced.

Poly(*p*-phenylene terephthalamide) PPD-T, can be polymerized to a fiber-forming molecular weight by polycondensation of terephthaloyl chloride and 1,4-phenylene-diamine.

p-phenylenediamine (PPD) — terephthaloyl chloride (TCL)

poly(*p*-phenyleneterephthalamide) (PPD-T)

This polymer could not be melted without decomposition, so the preferred fiber formation route was solution spinning. Patent literature suggests that the fiber is spun from a solvent system composed of dimethylformamide and lithium chloride. The final properties are achieved by stretching in steam after washing to remove residual solvent.

Poly(1,4-phenylene terephthalamide) (PPD-T) of high molecular weight can be prepared by low-temperature solution polymerization techniques. This polymer is less soluble in amide-type solvents than is poly(*p*-benzamide). The most successful conditions required hexamethylene phosphoramide (HMPA), the original solvent, alone or mixed with N-

methyl-2-pyrrolidone (NMP), although other mixtures such as NMP containing $CaCl_2$ also could produce a fiber-forming polymer. During polymerization, the molecular weight increases rapidly within the first few seconds of the reaction. The critical molecular weight or viscosity is exceeded, and the stir opalescence typical of lyotropic solutions is observed. Although gelation of the reaction mixture occurs quickly, polymerization continues, but at a greatly reduced rate. With the choice of a suitable solvent system, gelation can be delayed until the desired higher molecular weight is reached. Large-scale manufacture of a polymer requires continuous polymerization to minimize cost. In the case of PPD-T, special problems that had to be accommodated included the rapid gelation of the reaction solution with increasing molecular weight, the need to control the temperature of a vigorous exothermic reaction, and the handling of solvent HMPA, which is suspected to be a carcinogen.

Although melt spinning would be preferred from the standpoint of process simplicity and conversion cost, aramids must be spun from solutions, by wet, dry, or dry jet–wet methods because they decompose before or during melting, ruling out melt spinning. Dry spinning is used to produce Nomex® fiber where a dope (20%) of the polymer in solvent is converted to yarn. In wet spinning, the polymer dope is extruded into a nonsolvent where the fiber coagulates. The coagulated fiber then is washed and often drawn to develop desired fiber properties. In wet spinning the spinnerette is in the coagulation bath, hence the dope temperature and the coagulation temperature are the same. In dry jet–wet spinning (Fig. 21.25), as used for Kevlar®, on the other hand, the spinnerette is separated from the coagulation bath, allowing independent control of the dope extrusion and coagulation temperatures. The extrusion jet is placed a small distance above the coagulation bath, and the nascent fibers descend into the liquid, pass under a guide, and proceed in the bath while undergoing stretch; then they are withdrawn from the bath and wound up. A subsequent washing step may be required to remove residual acid solvent. In systems where the dope must be hot to lower viscosities to extrudable levels, and the bath cold for developing specific structures during coagulation, as is preferred with Kevlar®, dry jet–wet spinning is the option. For production of Kevlar® fibers, a PPD-T/concentrated sulfuric acid solution, containing approximately 20 percent polymer, is extruded at temperature less than 90°C into a cold water bath (~ 1°C). Following extrusion, the fiber is washed, dried, and given post-treatment, depending upon the properties desired.

Kevlar® is reported to have about twice the breaking strength (23–27 gram force/denier or 2.9–3.4 GPa) of high-tenacity nylon and polyester, but its most outstanding physical property is its high stiffness (550–1000 gram force/denier or 70–127 GPa), which can range up to an order of magnitude greater than that of standard polyester. This property has led to high-volume usage of the fiber as reinforcement in composite materials such as belts in radial tires and aerospace structures. The world production figure (1989) of this fiber was estimated to be over 70 million lb.

HIGH-TEMPERATURE-RESISTANT FIBERS

The need for high-temperature-resistance fibers has arisen from demands of a number of

Fig. 21.25. Schematic of dry jet–wet spinning method.

industrial applications, as well as applications in aerospace programs. In many of these applications, the usual characteristics of organic-based fibers are desired, but the high temperature resistance of inorganic fibers is required. Thus, the fibers are expected to retain their structural integrity at temperatures of 300°C and above for considerable periods of time, but otherwise their properties should resemble those of the more common manufactured textile fibers. The two leading groups of fibers in this area are the meta-aramids and PBI.

Meta-Aramid

The major fibers in this group, based on sales volume, are the products made from poly(*m*-phenylene-isophthalamide), which were introduced by Du Pont in 1962 as HT-1 nylon (later becoming known as Nomex®) and by Teijin in 1972 as Conex®. General manufacture of the fiber was described earlier under the heading "Aramids." Other manufacturers now are entering the field with products of similar chemical structure. Although these products are made and spun by different processes, their chemical and physical properties are similar. The fiber is usually utilized in the form of cut staple, which are amenable to conversion on traditional spinning and weaving machinery. Its mechanical properties (tenacity 4 to 5 gf/denier, ultimate strain 25 to 30 percent and initial modulus 90 to 100 gf/denier) compare favorably with those of other textile fibers. The main utility of the fiber lies in the resistance it offers to combustion: it has a limiting oxygen index (LOI) of about 0.29, a melting point above decomposition temperature, an ignition point above 600°C, and a flash point about 800°C.[18] The limiting oxygen index gives a relative measure of flame resistance; the higher the number, the lower the flamability. The fiber Kermel® from Rhone Poulenc also is classified as an aramid. It is, chemically, a polyamide-imide fiber and has an LOI of about 0.31. A major application of these fibers is in protective clothing. In order to reduce their cost, they often are blended with other, less expensive, flame-retardant fibers such as those based on cotton, rayon, and wool.

PBI

Poly-2-2′-(*m*-phenylene)-5,5′-bibenzimidazole, commonly called polybenzimidazole (PBI), was developed under the aegis of the U.S. Air Force Materials Laboratory in cooperation with the then-existing Celanese Corporation. The fiber went into commercial production in the United States in 1983. It is a condensation polymer obtained from the reaction of tetra-aminobiphenyl and diphenyl-isophthalate in a nitrogen atmosphere at temperatures that may reach 400°C in the final stages.[19] The structure of a repeating unit is shown below.

The polymer is dissolved at a high temperature under nitrogen pressure in dimethylacetamide, to which a small amount of lithium chloride may be added to increase the stability of the solution. Then, it is dry-spun in an atmosphere of heated nitrogen (about 200°C), from which the solvent is recovered; next it is stretched slightly in steam and washed. Drawing and relaxing are done in an inert atmosphere, as might be expected, because temperatures up to 250°C or higher are used. The fiber then is given a stabilization treatment in a sequence of steps and made into staple fibers using conventional crimping and cutting techniques. The stabilization treatment involves reaction with sulfuric acid and heating at high temperatures (~475°C) for short periods of time. The process, known as sulfonation, yields a product that has a significantly lower shrinkage than the unstabilized material.

The final yarn is golden yellow, and because this color appears to be an intrinsic property of the polymer, it may have some limitation as far as the civilian market is concerned. This material originally suffered from high

shrinkage on exposure to flame; however, further developments including the sulfonation treatment have reduced the shrinkage to only about 5 to 10 percent at 600°C.[20] The fiber is capable of retaining about one-half of its original strength (~3 gf/denier) upon exposure to air for 18 hours at 350°C or one hour at 425°C, it has an LOI of 0.41, which is well above that of the aramids, and it has high resistance to inorganic acids and bases and organic chemicals. Further, with tensile and moisture regain properties comparable to those of many textile fibers, the PBI fiber is well suited for blending with other fibers and conversion into final products using conventional spinning and weaving or nonwoven equipment. Applications include high-performance protective apparel, flight suits, and aircraft furnishings.

POLYTETRAFLUOROETHYLENE

Historical

Polytetrafluoroethylene (PTFE) was discovered in 1938 when Dr. Roy J. Plunkett and his assistants working on new nontoxic, nonflammable refrigerants at E. I. du Pont de Nemours and Co. found that one cylinder, which was supposed to contain tetrafluoroethylene (TFE), was empty. Upon opening the cylinder they discovered that the inside was covered with a white powder. Polymerization of tetrafluoroethylene had taken place. The result was polytetrafluoroethylene (PTFE) or Teflon®, which is the trade name applied to that polymer by du Pont. The company tested the polymer and found that it was virtually inert to all known solvents, acids, and bases, a characteristic that was unique at that time. It also found that this material was resistant to high temperature and had the lowest coefficient of friction of any known solid. The research on PTFE was intensified during World War II because of military demands for improved materials for products such as gaskets, packing, and linings for containers for handling corrosive materials. When the war was over, the release of classified information caused a booming interest in the polymer. Since that time, PTFE has found itself in many different applications.

Manufacture

In one process, the manufacture of the monomer, tetrafluoroethylene, involves the following reactions: Hydrogen fluoride is made by reacting calcium fluoride with sulfuric acid:

$$CaF_2 + H_2SO_4 \rightarrow CaSO_4 + 2HF$$

Chloroform is reacted with hydrogen fluoride in the presence of antimony trifluoride as the catalyst:

$$CHCl_3 + 2HF \rightarrow CHClF_2 + 2HCl$$

Tetrafluoroethylene is obtained by the thermal decomposition of this monochlorodifluoromethane (known as Freon) in a continuous noncatalytic gas-phase reaction, carried out at or below atmospheric pressure at temperatures from 600°C to 900°C:

$$2CHClF_2 \rightarrow C_2F_4 + 2HCl$$

Numerous side-products are generated in this process. Many of them are present in trace amounts, but the highly toxic perfluoroisobutylene, $CF_2 = C(CF_3)_2$, requires special precautions.

TFE also may be manufactured by the reaction of zinc and tetrafluorodichloroethane:

$$ClF_2C{-}CF_2Cl + Zn \rightarrow C_2F_4 + ZnCl_2$$

or by the reaction of tetrafluoromethane molecules in an electric arc furnace:

$$2CF_4 \rightarrow C_2F_4 + 2F_2$$

Tetrafluoroethylene is a colorless, tasteless, odorless, and nontoxic gas. To avoid any undesired reactions during storage, inhibitors must be added. The polymerization is carried out by an addition-type reaction in an aqueous emulsion medium and in the presence of initiators such as benzoyl peroxide, hydrogen peroxide, and persulfates. The monomer is fed into a cooled emulsion medium, and then heated to a temperature of 70 to 80°C, at which the polymerization takes place. The

pressure may range from 40 to 100 atmospheres. After removal of the unreacted material, the polymer is washed, pressed, and dried. PTFE is manufactured in four different forms: granular, fine powder, aqueous dispersion, and micro powder, and in a variety of grades, each differing in properties. Fillers such as glass fibers, asbestos, graphite, or powdered metals may be added to the granules in order to modify properties.

Because Teflon® is not soluble, it cannot be wet- or dry-spun, and because it is thermally unstable at its melting point of about 400°C, this combination would seem to pose an impossible problem for the production of fibers. Research into the fundamental characteristics of the polymer, however, revealed that the submicroscopic particles precipitated from the polymerization reaction were about 100 times as long as they were thick.

In one manufacturing process, an aqueous dispersion of PTFE is mixed with a solution of ripened cellulose xanthate, from which a fiber is obtained by a wet-spinning process, after which the cellulose is completely decomposed by heating. The remaining PTFE is sintered into continuous fibers by transporting them over heated metal rolls, followed by stretching to achieve the desired diameters and physical structure. In another process, films are produced that are slit into strips of very small widths, which then are stretched and sintered. In yet another, tapes are extruded, which are converted to fibers by stretching while being twisted to a very small cross-sectional area.

The chemical inertness and the thermal stability of this polymer are so great that in spite of its high price ($4–$10/lb, depending on the resin type) it is used in chemical operations where drastic conditions exist and no other organic material is suitable.

GLASS AND CARBON FIBERS

Glass

Among the manufactured inorganic fibers, glass is produced in by far the largest volume. There has been a rapid increase in the use of textile grades of these fibers, and outside the textile field enormous quantities of glass fibers are used in air filters, in thermal insulation (glass wool), and for the reinforcement of plastics.

Glass possesses obvious and well-known characteristics which have largely determined the methods used to form it into large objects. It flows readily when molten and can be drawn into filaments, whose extreme fineness appears to be limited only by the drawing speed. The method used in producing textile-grade glass fibers follows this principle. (See Fig. 21.26.)

In the commercial operation the molten glass, produced either directly from raw materials or by remelting of marbles, is held at a uniform temperature in a vessel, whose bottom carries a bushing containing small uniform holes. The molten glass flows through these holes as tiny streams that are attenuated into filaments at speeds on the order of 3000 meters/min; these flaments are coated with a lubricant, gathered into groups to form yarns, and wound up. For a particular glass viscosity, the size of the individual filaments is determined by the combination of the hole size and the speed of attenuation.

Because of the inherently high modulus of glass, very fine filaments are required in order to approach the required properties of textile materials. Thus, the diameter of glass filaments falls in the range of 3.8 to 7.6 microns, whereas the average diameter of the finest organic fibers is about twice as great. The fiber and yarn numbering system is based on nomenclature used in the glass industry and differs from the traditional systems accepted in the textile and organic fiber industries.

The method of manufacture of glass staple fibers differs from those used to produce the corresponding organic materials, all of which are based upon cutting the continuous filament product. Air jets, directed in the same line of flow as the emerging streams of glass, attenuate the streams, and break the solid glass into the lengths desired for further processing, which are gathered on an appropriate vacuum drum and delivered as slivers or a matte. To produce fibers that may be

Fig. 21.26. Flow diagram for the manufacture of textile glass fiber: (1) glass batch; (2) batch cans; (3) marble-forming; (4) cullet cans; (5) marbles; (6) melting furnaces; (7) filament yarn formation; (8) gathering and sizing; (9) yarn packaging; (10) air jets; (11) lubricant spray; (12) collection for staple fibers; (13) staple fiber packaging. (*Courtesy Owens-Corning Fiberglass Corp.*)

coarser and considerably less uniform in length, to be used for the production of filters, paper, or thermal insulation, large streams of molten glass are cross-blown by blasts of hot air, steam, or burning gas.

As might be expected from the nature of glass, the conversion of glass fibers into final products has required the development of new lubricants, finishes, and processing techniques. For example, because glass fabrics cannot be dyed directly or printed with the colors demanded for their acceptance as draperies, the colorant must be applied to a resin coating. But before applying the coating it is necessary to remove the lubricant that was placed on the fibers to permit their conversion into a fabric, which is done by burning. The elevated temperature resulting from this operation also relaxes the internal strains developed in the glass fibers during the steps of the textile operations and sets the yarns in the required geometry. The fabric then is resin-treated, cured, and dyed or printed.

Another inherent property of glass is the tendency of unprotected fiber surfaces to abrade each other to destruction under the action of very little mechanical working. When it was first considered for rubber-reinforcing purposes, its poor adhesion to rubber and the inadequacy of the then-existing bonding agents frustrated attempts by manufacturers to take advantage of the very high tensile strength, completely elastic behavior, high modulus, and lack of moisture sensitivity of glass fibers. However, it has been possible to modify the fiber surfaces so that satisfactory adhesion is achievable, and the impregnant can be applied in such a way that fiber-to-fiber contact is prevented. With the adhesion problems solved, glass in cord form could effectively enter markets in belt-type tire construction and in all kinds of power-transmitting rubber belts. It is estimated that the total glass fiber production in the United States in 1990 was approximately 1.5 billion lb.

Carbon and Graphite

Following World War II the development of jet aircraft and rockets created demands for fibers having thermal resistance, strength, and modulus far beyond what could be obtained in existing organic fibers. Much of this need was for reinforcing materials that could be embedded in matrices of one type or another. As a result, techniques have been developed for preparing fibers from a good many metals

and refractory inorganic compounds. Although these materials are essential for certain uses, the volume of production still is low, and the prices are correspondigly high (as much as $1,000/lb).

Carbon and graphite fibers are made from rayon and acrylic precursors by driving off virtually all of the hydrogen and oxygen contained in them. The principle is essentially the same as that which brought about the formation of coal, or, citing a more recently discovered and dramatic example, the conversion of the original wooden beams of buildings in Herculaneum, buried by a flow of mud from Vesuvius in A.D. 67, to what appears to be charcoal. In the present commercial process the starting material is selected so as to produce a final product of the desired size and properties. Cotton, bamboo, and other natural fibers were the earliest materials used as precursors. In the 1950s, rayon was used for this purpose, but the first attempts led to fibers with poor strength. Union Carbide, in the mid-1960s, made the first strong carbon fiber from rayon and extended its work to using polyacrylonitrile as the pecursor. The use of mesophase pitch as the starting material began in the 1970s. Carbon fibers may be divided into three types: (1) highly graphitized with a high modulus, (2) heat-treated at a lower temperature to produce fibers with lower modulus but high strength, and (3) randomly oriented crystallites with relatively low modulus, and low strength, but most important, low cost.

The preparation of fibers generally consists of heat-treating the precursor at a low temperature (200–350°C), usually in air, which gives a stable fiber for higher-temperature processing. During this step extensive decomposition occurs, and a percentage of the initial weight is lost, which is related to the cleavage of the C—C and C=O bonds and expulsion of H_2O, CO, and CO_2 (additionally HCN in the case of PAN precursor) as gases. This step is followed by carbonization in an inert atmosphere at 1000 to 2000°C, which is said to collapse the cyclized structure into a stacked ring carbon fiber structure. Almost all noncarbon elements are evolved as volatiles. This is followed by graphitization, which usually is carried out at temperatures above 2500°C for short periods in argon or nitrogen. The process increases the purity, removes the defects, and further improves the order in the structure. Thus, the difference between the so-called carbon and graphite fibers lies in differences in the ranges of temperatures at which the last step is carried out, the degree of carbon content (97% for carbon and 99.6% for graphite), and the mechanical properties, which are superior in the graphite fibers. In usual discussions the term "carbon fibers" covers both materials. The majority of carbon fibers produced today are made from a PAN precursor. PAN fibers are fine, and have a higher degree of molecular order and a higher decomposition point than those from rayon. This precursor also leads to a greater carbon yield (45%, as opposed to 24% from rayon), but the fibers are more expensive to produce than those based on rayon. Recent commercial developments have allowed the production of carbon fibers from low-cost petroleum or coal-tar pitch, instead of synthetic fibers. These precursors lead to a higher yield (90%) of carbon, improved lubricity of fiber products, and higher production rates. However, the pitch-based fibers may be more brittle and harder to handle, have a higher specific gravity (2.0, as opposed to 1.8 for those from PAN and 1.66 from rayon) and lower compatibility with some matrix materials.

The fibers have a diversity of applications. One major application is in composites, where they are used for reinforcing resins and metals to provide structural materials with high strength, high modulus, and light weight. The resulting composites are used in the aircraft and aerospace (the largest users), automotive, and sports industries. The fibers also find uses in protective garments, electrical devices, insulation, and filtration. The use of the fibers in the United States increased from about 1 million lb in 1980 to over 5 million lb in 1990. The prices dropped significantly during that decade, but in 1990 the majority remained in the $15 to $70/lb range. The price of the ultra-high-modulus carbon fibers, however, can be as high as $1500/lb, or greater.

SULFAR

Historical

Sulfar fibers are extruded from poly(phenylene sulfide) or PPS by the melt-spinning process. The first PPS polymer was made in 1897 by the Friedel-Crafts reaction of sulfur and benzene. Researchers at Dow Chemical, in the early 1950s, succeeded in producing high-molecular-weight linear PPS by means of the Ullmann condensation of alkali metal salts of *p*-bromothiophenol.

In 1973, Phillips Petroleum Company introduced linear and branched products under the trade name Ryton® by reacting 1,4-dichlorobezene with sodium sulfide in a dipolar aprotic solvent. In 1983, the same company succeeded in stable melt-spinning of PPS. In 1986, the Federal Trade Commission gave the fiber the generic name Sulfar, defined as "a manufactured fiber in which the fiber forming substance is a long chain synthetic polysulfide in which at least 85% of the sulfide (—S—) linkages are attached directly to two aromatic rings."

Manufacture

In one process, synthesis was carried out by self-condensation of a metal salt of a *p*-halothiophenol:

where X is a halogen, preferably Br, and M is a metal (Na, Li, K, or Cu, preferably Cu). This reaction was carried out under nitrogen in the solid state or in the presence of materials such as pyridine as reaction media. Considerable difficulty was encountered in removing the by-product, copper bromide, from polymers made by this process. The current commercial synthesis of the polymer is carried out by reaction between *p*-dichlorobenzene and sodium sulfide in a polar solvent.

The process discovered by workers in the laboratories of Phillips Petroleum Co. marked a significant departure from prior processes, and made it possible to prepare a variety of arylene sulfide polymers from the readily available starting materials.[21]

Melt spinning of PPS involves problems such as plugging of the filter and the spinnerette. In order to prevent cross-linking and gel formation during the process of spinning, the company introduced cure retarders comprising Group IIA or Group IIB metal salts of fatty acids, which improved the heat stability of PPS:

$$(CH_3(CH_2)_nCOO)_2M$$

where M is a Group IIA or IIB metal, and n is an integer from 8 to 18. Representative compounds of the type described above include calcium stearate, calcium laurate, calcium caparate, and calcium palmitate. In general, the cure-retardant additives are employed in an amount within the range of about 0.1 to 5, preferably about 0.5 to 2, weight percent based on the weight of the PPS.

Sulfar fibers, sold under the trade name Ryton® in the United States, are characterized by high heat resistance, inherent flame retardancy, excellent chemical resistance, low friction coefficient, good abrasion resistance, and good electrical properties.[22] Physical characteristics include medium tenacity (3.5 gram force/denier or 423 MPa) and elongation (25–40%) and low shrinkage (<5% @ 100°C). The fibers find application in a number of industrial products, including filter fabric for coal-fired boiler bag houses, paper maker's felts, materials for electrical insulation, high-performance composites, gaskets, and packings.

Several Japanese and European companies have begun the production of PPS, some with a U.S. partner. The 1990s should see the

introduction of a number of new fibers and fiber products based on the PPS polymer.

MICRODENIER FIBERS

One of the more important developments in the field of fiber technology in recent years has been the production of fine denier fibers, with worldwide activity in developing new products and outlets for them. Improvements in the quality of polymers, coupled with new technology for extrusion of fibers, have led to the production of fibers with sizes ranging from 1 denier to 0.1 denier or even lower. Fabrics produced from such fibers have novel and unique properties, and they are finding applications in a wide variety of apparel, and industrial products, including high-fashion fabrics with silk-like texture, synthetic suede, breathable porous but rainproof or bacteria-proof fabrics, wipes for oil and other spills, clean room materials, and cloths for cleaning camera, microscope, and spectacle lenses and for compact discs.

Unfortunately, there is no universally accepted definition for micro denier, and companies have been free to use terms they choose. In general there seems to be a consensus that the term "fine" may be used for a denier of 1 or less and "micro" for a denier less than 0.5. The production of such fibers is difficult and expensive because the throughput rate must be reduced in order to obtain fine denier filaments, and there is a limit to how fine in denier a regular process could be made to go. Generally, the technology of production of fine and microdenier filaments can be divided into five categories. First, conventional spinning technology can be used, by using fine-size dies and adjusting the throughput rate and quenching and drawing parameters to obtain fine denier polyester filaments. This method has been used by most fiber manufacturers to produce such fibers in limited quantities. In another method, alkali reduction or surface etching is used to dissolve the surface layers, in an effort to reduce the cross-sectional size of filaments in a polyester fabric. A weight loss of as much as 25 percent has been achieved by this process. In the melt-blown process, molten polymer is forced through a melt blowing die and die tip orifices, and the emerging stream is attenuated by a jet of high-velocity hot air until broken. Then the broken fiber is forced into a stream of cold air, where the fiber solidifies. The fiber is collected on a wire screen or apron with other fibers, and a homogeneous melt-blown web is produced. This method is used extensively with polypropylene materials, but other polymers such as nylon, polyester, and polyethylene also have been used. Fibers as fine as 0.5 micron and finer have been produced.

In another approach, filaments containing two polymers that do not adhere to each other are spun and then split. One may, for example, spin a bicomponent fiber of nylon that has several filaments of polyester embedded. After a fabric containing bicomponent filaments has been woven, it is treated to split the components, thus converting the original filament to several smaller filaments. Deniers on the order of 0.1 can be achieved by this ingenious method. In the last method, instead of splitting the two components as in the previous example, one component is dissolved away chemically, leaving bundles of very fine fibers in the fabric.

The majority of the technology for producing fine and micro denier fibers is new and thus expensive; so efforts during the 1990s can be expected to be directed toward optimizing process parameters in existing methods and discovering faster and cheaper methods for manufacturing these aesthetically very pleasing and functionally very promising fibers.

FIBER VARIANTS

Introduction

In a previous section, data and plots were given showing the rapid rise in consumption and production of manufactured fibers at the expense of natural fibers. The principal reason for this has been the wide range of manufactured fiber variants that can be produced from a single fiber-forming polymer. The wide range of polymers available, each with its particular

properties, adds yet another dimension. This is not to say that there is only one type of cotton, wool, silk, or asbestos fiber; there are many varieties of natural fibers, but their supply is limited by natural factors such as climate and genetics. The relative availabilities of manufactured fiber types can be altered by controlled chemical-process changes, whereas the amount and the quality of a desired cotton type that can be grown is determined to a great extent by climatic conditions, which humans have not yet learned to control. Another factor that has aided the growth of manufactured fibers is their consistent quality and properties. Again the grade and the quality of natural fibers are subject to the vagaries of nature.

For the purposes of this discussion, fiber variants will be divided into two types: chemical and physical. Chemical variants will be those involving a small but significant change in composition, whereas physical changes will be those involving a change in either the dimensions of the fiber or its stress/strain or stability features. The definitions of the two variants also could be based on modification of either aesthetics or functionality.

Physical Variants

Most manufactured fibers are available as staple, tow, and filament. Natural fibers are available only in the characteristic forms in which they occur, with filament silk and cotton staple as examples. All manufactured fibers are formed initially as filament yarns. The German adjective *endlos* (literally translated "endless") is very descriptive, as filament yarns are continuous strands consisting of one or more members that for most practical purposes are infinite in length. Fine filament yarns (40 to 100 denier) are used in producing lightweight apparel fabrics, whereas coarse filament yarns (800 to 1200 denier) are found as reinforcement in tires or conveyor belts. These examples are chosen as extremes to show the range or applicability of manufactured fibers, and represent only a small fraction of the actual range of end uses.

When many filament yarns are collected into a bundle immediately after formation, the resulting structure is called a tow. Tows may range from 10,000 to over a million in total denier. In a next step, the tow may be crimped by the process previously described, which imparts what is usually a saw-tooth appearance along the length of all the filaments. For some end uses, the crimped tow itself is provided by the fiber maker, as are, for example, the previously discussed acrylic tows, which are converted to staple as part of the spinning of yarns with a wool-like character. Another example is the cellulose acetate tows used to form cigarette filters or the ink reservoirs for marker pens. In the latter case, the compact tow bundle first is treated to separate the individual filaments, giving a voluminous structure, which then is gathered into a continuous rod, wrapped in paper, and cut into appropriate lengths.

In the fiber-making plant, tow may be cut into short lengths of staple, ranging in length from $\frac{1}{16}$ inch to 6 inches, depending upon the end use. For instance, the very short staples are used either in making flocked structures or in the production of papers containing blends of natural cellulose and manufactured fiber. The longest lengths are used in spinning heavy yarns for carpets or cordage. But most staple is $1\frac{1}{2}$ to 3 inches in length and is used to form blends with cotton, rayon, or wool in the yarns employed in standard apparel fabrics. The staple length of the manufactured fiber is chosen to match that of the other blend component; otherwise, uneven yarn of poor quality results.

The size of a manufactured fiber can be altered by changing the size of the hole through which it is extruded and maintaining a constant takeup speed. Commercial fibers range from about 1.25 denier to 25 denier, corresponding to average diameters of about 5 to 50 microns. The very large ones would be used to make doll's hair or wigs, but the majority of manufactured fiber staples are made in the 1.5 to 6.0 denier range, corresponding to cotton blending fiber at the low denier end and coarse wool blending fiber at the high end. Staple and filament yarns used in carpets are in the 12 to 16 denier

range, whereas industrial filament yarns such as tire cord are about 6 denier per filament. The size of a fiber is a determinant of its stiffness, which in turn influences the draping, quality and the surface feel (often called hand in the textile industry) of a fabric made from it.

Although fibers normally are extruded through circular jet holes, the use of noncircular holes in the jets has led to the availability of a wide variety of cross-sectional shapes. In the case of fibers dry- or wet-spun from solutions, most of the mass exiting from the jet hole is not the polymer. For example, cellulose acetate fiber is made from a 25 percent solution of cellulose acetate in acetone containing a small amount of water. After the fiber leaves the jet face, solvent begins to evaporate, and as a result the area of the fiber cross section decreases. The final result is a fiber of roughly circular cross section, but with a serrated edge, and much smaller in area than the parent jet hole. When cellulose acetate is extruded through a triangular jet hole, the end result is a fiber of "Y" cross section, due to shrinkage from the original triangular shape.

In the melt-spinning process, there is no solvent loss to influence final cross-sectional shape. In the case of a triangular jet hole in melt spinning, the molten fiber leaves the jet face with a triangular cross section but, being molten, immediately tends to return to a circular cross section due to surface tension forces. It is necessary, therefore, to quench or cool the fiber as soon as possible in order to maintain the desired cross-sectional definition. Some loss of this definition is unavoidable during the drawing step that normally follows extrusion in a melt-spinning process.

Fibers of noncircular cross section can modify and change both functional and aesthetic properties in textile structures. The triangular cross section is typical in those respects; its shape leads to a stiffer fiber than circular fiber of the same cross-sectional area, and in a fabric this results in less drapability and a crisper surface feel. Also, the flat surfaces reflect light in a different way than do curved surfaces and can create desirable lustrous effects. These optical effects are subject to many subtle influences having to do with the size of the reflecting surface and the amount of internal reflection that takes place. Triangular or "Y" cross-sectional fibers have greater specific surface areas per unit weight than their circular counterparts, which accounts for their use in aerosol filtration, where surface area is a major factor in efficiency. Yarns made with triangular cross-section fibers are more voluminous than those from round-cross-section fibers. Thus when the two types are made into fabrics of equal weight, the variant cross-sectional fabric will transmit less light and be less permeable to air than that made with round-cross-section fibers. In the spinning of blended staple yarns, maximum strength is obtained if the blend components have similar load/elongation characteristics. The stiffness of polyester staple fibers can be varied mainly by changing the draw ratio used to orient the fiber after extrusion and thus producing fibers suitable for blending with cotton, rayon, acrylic, or wool fibers, which possess markedly different properties. In many industrial applications, fiber and yarns are used under conditions where they bear a load while hot, as for example in tires and power transmission belts, where it is important that they do not grow or stretch significantly under these conditions. Accordingly, industrial yarns are drawn to a greater extent than other yarns to reduce as much as possible the stretch of the final products. Relatively, these yarns have high strengths, high stiffnesses, and low elongations at rupture.

Although in most instances it is desirable to have fibers that are dimensionally stable, in some structures it is advantageous to use mixtures of fibers that are stable with fibers that do shrink upon exposure to heat or steam. High shrinkage potential usually is built into a fiber by stretching it and not giving it a stabilization treatment. The yarn bulking that occurs with blend yarns of high- and low-shrinkage acrylic staples has already been described, in the section on that fiber. Using the same principle, felt-like structures can be made by heat treatment of nonwoven battings containing low and high-shrink staple.

Chemical Variants

The use of titanium dioxide as an additive in the delustering or dulling of manufactured fibers was discussed in the section on viscose rayon. The addition of pigment particles influences the processing and performance of the fiber, along with changing its appearance. Because of the whiteness of the delustered fiber, it requires more dyestuff to reach a given shade than that needed in the case of a bright fiber. The sliding friction of a delustered fiber is lower because the pigment particles protruding from the fiber surface reduce the contact area between a fiber and the surface it slides against—a guide, for example. By the same token, some pigments can accelerate the wear of contacted surfaces. It has been suggested that the drawing and consequent orientation of delustered fibers proceeds more smoothly because the pigment particles act as nucleation sites where molecular motion is initiated. Unless specially treated, the surface of anatase—one of the crystalline forms of titanium dioxide—can accelerate the ultraviolet light degradation of acetate or nylon fibers. It is postulated that the crystal surface catalyzes the formation of peroxides from the water and oxygen under the influence of ultraviolet light, and that peroxides are the active species in the resulting polymer degradation. For this reason, rutile—the other common crystalline form of titanium dioxide—is used to deluster fibers when improved sunlight resistance is needed.

The degree of polymerization or the molecular weight of the polymers on which manufactured fibers are based can be controlled as part of one of the early steps in the process. The polymer molecular weight chosen for a fiber has a strong influence on process economics, ease of conversion to fiber, and end-use performance. Commercially acceptable fibers are based on the best balance of these factors. However, in fabrics that are open in texture and made from standard polyester staple spun yarns of low twist, a condition known as "pilling" will develop as a result of wearing. The pills are made up of fiber ends that have worked loose from the yarn bundles as a result of surface rubbing and have wrapped around themselves. In the case of fabrics from natural fibers, which are generally less wear-resistant than manufactured fibers, these pills or fiber bundles will be lost by attrition with continuing wear. Because the wear resistance of a manufactured fiber is related to its molecular weight in a general way, the pilling tendency of a polyester staple can be reduced by lowering its degree of polymerization. This compromises the tensile characteristics of the fiber only to a small extent, and all the other desirable properties such as minimum care characteristics are essentially unaffected. When the molecular weight of a polyester or a nylon polymer is increased above the standard, the resultant fibers will have increased tensile properties and fatigue resistance. The filament polyester and the nylon yarns used in end products such as ties and conveyor belts are based on such polymers.

The technology for "dope dyeing" or mass coloration of fiber as part of the fiber manufacturing operation was described in the sections on viscose rayon and cellulose acetate. The use of this technology has decreased, largely because of problems of profitably managing the required inventory of colors in rapidly changing fashion markets, and for this reason it is used extensively only if the fiber cannot be dyed by any other means. Polypropylene fiber for use in outdoor carpets is a good example of such a situation. However, if one regards "white" as a color, one finds that a substantial portion of the polyester staple fiber produced for blending with cotton or rayon contains an optical brightener or a fluorescing agent. This is needed to overcome the yellowing tendency of polyester following the absorption of hydrophobic soils. The cotton or rayon fibers in a blend are continually rewhitened by the fluorescing agents added to laundry detergents for that purpose, but these agents are without effect on the polyester blend component. Because the polyester component usually is at least half of the blend, the spun-in optical brightener it contains is vital for the maintenance of overall whiteness.

In the cases of polyester, nylon, and acrylic

fibers, their manufacturers have developed fiber variants with a wide range of dyeing behavior, referred to as dye variant fibers. Polyester fibers usually are dyed with what are described as disperse dyes. These dyes are only slightly soluble in boiling water and are used in the form of dispersions. The dye in aqueous solution is assumed to be in monomolecular form, and is absorbed from the dyebath into the polyester fiber by a process often called solid solutioning. As this occurs, more solid dye is dissolved to replenish that entering the fiber. The overall rate of dyeing is a very complex phenomenon but in part is determined by the molecular structure of the fiber through which the dye must diffuse. Generally, the more that the polymer molecules have been organized by drawing and annealing into more geometrically perfect domains or crystallites, the slower is the dyeing rate. The ability of the molecules to be thus organized into compact ordered structures can be reduced by polymerizing a small amount of a foreign dibasic acid or glycol into the polymer. Usually, 5 to 10 mol% is sufficient to prevent this regularity. Adipic acid, isophthalic acid, and polyethylene glycols are used to produce some of the comonomers for fast- or deep-dyeing polyesters. In this way, these fibers are made more economical to dye or print because special dyebath additives, high dyeing temperatures, and high-pressure steam-print fixation usually are not required.

Polyester fibers can be given an additional mechanism for dyeing if an ionic comonomer is added during polymerization. A common additive is an alkali metal salt of dimethyl-5-sulfo-isophthalate, which gives sulfonic (anionic) groups as part of the polymer structure. These groups allow the fiber to absorb basic (cationic) dyes by a specific ionic mechanism. The amount of cationic dye that can be absorbed by the fiber is stoichiometrically related to the number of anionic sites present in the fiber; this is quite distinct from the general solid solutioning that takes place with disperse dyes and polyester. A cationic dyeable polyester is useful for two main reasons in fabric coloration. First, cationic dyes give brighter, clearer shades than disperse dyes, and this can be important for both solid-dyed shades and prints during ever changing fashion cycles. Second, fabrics containing arrangements of unmodified and cationic dyeable polyesters can be dyed in the piece to a variety of color/white combinations by selection of dyestuffs, as day-to-day changes in demand may require. This is more economical than dyeing yarns to different colors, holding them in inventory, and then weaving them into fabrics, if and when they are required.

Nylon 6,6 can be dyed either with disperse dyes or with acid (anionic) dyes, the former carried out by essentially the same steps as described above for polyester, except that nylon absorbs these dyes much more readily. The ability of nylon to absorb acid (anionic) dyes is the result of a significant number of accessible free amine (cationic) end groups being present in the polymer. The dyeing of nylon with acid dyes is analogous to the dyeing of wool or to the dyeing of modified polyester with basic dyes, except that the polarities of the interacting groups are reversed in the latter case.

For best fastness to light and washing, nylon 6,6 is dyed with acid dyes, and nylon dye variants thus are based on manipulation of the level of acid dye uptake. By adding a monobasic acid such as acetic acid to the reaction mix near the end of the polymerization process, the amino end groups are converted to amide groups which have no affinity for acid dyes under normal dyeing conditions. This technique creates light acid dyeing or acid reserve dye variants that can have some capacity to absorb basic dyes at the carboxyl end groups. Nylon dye variants with increased acid dye uptake can be made by using a slight excess of diamine in the polymerization. In this way, there are no free carboxylic acid end groups. Nylon dye variants have found the greatest acceptance in floor coverings where attractive patterns can be piece-dyed using controlled dyebath conditions and selected acid dyestuffs.

Acrylic fibers are dyed most frequently with basic dyes. This is made possible by

copolymerizing acrylonitrile with an acidic monomer such as styrene-*para*-sulfonic acid. Acrylic fiber suitable for acid dyeing can be made by using a basic comonomer such as a vinyl pyridine or a vinyl pyrrolidone.

Fibers not having inherent flame resistance often can be given this property by incorporation of a suitable additive. This may be done by copolymerization of the additive into the polymer reaction of the additive with the polymer after polymerization, or by applying a polymeric or monomeric noninflammable finish to the surface as a coating. These additives usually contain bromine, nitrogen, or phosphorus or a combination of these elements. Great care must be taken in choosing the additive and its level of addition in order to prevent loss of other desirable fiber properties and to avoid any harmful effects to processors or ultimate consumers.

A wide variety of special durable surface treatments have been used on manufactured fibers. These include treatments for imparting such characteristics as soil resistance, antistatic behavior, and wearer comfort through moisture wicking and transport. Fiber finishes also have been used successfully in promoting adhesion between two materials, as, for example, between polyester tire cord and rubber, and between glass fiber and polyester resin.

REFERENCES

1. Time-Life Books, *Seven Centuries of Art*, New York, 1970.
2. Mark, H., and Whitby, G. S. *Collected Papers of W. H. Carothers*, John Wiley & Sons, New York, 1940.
3. Turbak, A., "Rayon," in *Encyclopedia of Polymer Science and Engineering*, 2nd ed., Vol. 14, p. 55, John Wiley & Sons, New York, 1985.
4. Davis, S., *Textile Horizons*, **9** (2), 62 (1989).
5. Markham, J. W., *Competition in the Rayon Industry*, p. 16, Harvard University Press, Cambridge, MA, 1952.
6. Robinson, J. S., *Fiber-Forming Polymers: Recent Advances*, Noyes Data Corp., Park Ridge, NJ, 1980.
7. Barnes, C. E., "Nylon 4—Development and Commercialization," *Lenzinger Berichte*, No. 62: 62–66, March 1987.
8. O'Sullivan, D., *Chemical and Engineering News*, **62** (21), 33 (1984).
9. Davis, G. W., Everage, A. E., and Talbot, J. R., *Fiber Producer*, **12** (6), 45 (1984).
10. Ahmed, M., "Polypropylene Fibers—Science and Technology," in *Textile Science and Technology 5*, p. 16, Elsevier, New York, 1982.
11. Lieberman, R. B., and Barbe, P. C., "Propylene Polymers," in *Concise Encyclopedia of Polymer Science and Engineering*, J. I. Kroschwitz (Ed.), p. 916, John Wiley & Sons, New York, 1990.
12. Hogan, J. P., and Banks, R. L., "History of Crystalline Polypropylene," in *History of Polyolefins*, R. B. Seymour and T. Cheng (Ed.), p. 103, D. Reidel Pub. Co., Boston, 1986.
13. Zwijnenburg, A., and Pennings, A. J., *Colloid and Polymer Science*, **259**, 868 (1978).
14. Smith, P., and Lemstra, P. J., *J. Materials Science*, **15**, 505 (1980).
15. Kavesh, S., and Prevorsek, D., U.S. Patent 4,413,110, to Allied Chemical, 1983.
16. Kwolek, D. L., U.S. Patent 3,600,350, to E. I. du Pont de Nemours and Co., 1971.
17. Blades, H., U.S. Patent 3,767,756, to E. I. du Pont de Nemours and Co., Inc., 1973.
18. McIntyre, E., *Textile Horizons*, **8** (10), 43 (1988).
19. Chenevey, E. C., and Conciatori, A. B., U.S. Patent 3,549,603, to Celanese Corp., 1970.
20. Coffin, D. R., Serad, G. A., Hicks, H. L., and Montgomery, R. T., *Textile Research Journal*, **52**, 466 (1982).
21. Edmonds, J. T., Jr. and Hill, H. W., Jr., U.S. Patent 3,354,129, to Phillips Petroleum Co., 1967.
22. Scruggs, J. G., and Reed, J. O., "Polyphenylene Sulfide Fibers," in *High Technology Fibers, Part A*, M. Lewin and J. Preston (Eds.), Marcel Dekker, Inc., New York, 1985.

SUGGESTED READING

The reader is referred to the four encyclopedias listed below for additional information. They contain enormous quantities of information on manufactured fibers as well as comprehensive bibliographies.

Concise Encyclopedia of Polymer Science and Engineering, John Wiley & Sons, New York, 1990.

Encyclopedia of Polymer Science and Engineering, 2nd ed., John Wiley & Sons, New York, 1985. (Seventeen volumes, index volume, and supplement volume.)

Encyclopedia of Polymer Science and Technology, Interscience Publishers, New York. (Sixteen volumes.)

Kirk-Othmer Encyclopedia of Chemical Technology, 3rd ed., Interscience Publishers, New York. (Twenty-one volumes and a supplement, 3rd ed.; to date, 16 volumes.)

The following books contain broad discussions of manufactured textile fibers.

Baer, E. and Moet, A. (Eds.), *High Performance Polymers*, Hanser Publishers, New York, 1991.

Billmeyer, F. W., *Textbook of Polymer Science*, John Wiley & Sons, New York, 1984.

Ciferri, A., and Ward, I. M. (Eds.), *Ultra-High Modulus Polymers*, Applied Science Publishers, London, 1979.

Datye, K. V., *Chemical Processing of Synthetic Fibers and Blends*, John Wiley & Sons, New York, 1984.

Hearle, J. W. S., and Peters, R. H. (Eds.), *Fibre Structure*, The Textile Institute, Manchester, Butterworths, London, 1963.

Lewin, M., and Preston, J. (Eds.), *Handbook of Fiber Science and Technology: High Technology Fibers*, Vol. III, Marcel Dekker, Inc., New York, 1983.

Mark, H. F., Atlas, S. M., and Cernia, E. (Eds.), *Man-Made Fibers; Science and Technology*, Vols. I, II, and III, John Wiley & Sons, New York, 1967, 1968, and 1968.

Moncrieff, R. W., *Man-Made Fibres*, 6th ed., John Wiley & Sons, New York, 1975.

Morton, W. E., and Hearle, J. W. S., *Physical Properties of Textile Fibres*, The Textile Institute, Manchester, Butterworths, London, 1975.

Peters, R. H., *Textile Chemistry*; Vol. I, *The Chemistry of Fibers*, and Vol. II, *Impurities in Fibers; Purification of Fibers*, Elsevier, New York, 1963 and 1967.

22

Synthetic Organic Chemicals

Daniel E. Collins* and Forrest A. Richey, Jr.**

Synthetic organic chemicals can be defined as products derived from naturally occurring materials (petroleum, natural gas, and coal), which have undergone at least one chemical reaction, such as oxidation, hydrogenation, or sulfonation.

The volume of synthetic organic chemicals produced in the United States increased from about 42 billion pounds in 1958 to more than 270 billion pounds in 1990. The growth in production for the past 35 years is shown in Fig. 22.1. The effect of the economic slowdowns in 1974/75 and in the early 1980s on chemical output is reflected very clearly. Although the total volume increased from 1989 to 1990, the sales value actually decreased, presaging the slowdown of the early 1990s. Up until the mid 1960s, most of the phenomenal growth reflected the replacement of "natural" organic chemicals. Since that time, growth for synthetic materials has been dictated by the expansion of present markets and the development of new organic chemical end uses. It is not certain that this rate of growth can be maintained as the uses of many products become strongly affected by environmental concerns. In fact, the future may see more emphasis on chemicals that can be produced from renewable raw material sources.

Those synthetic organic chemicals (excluding polymers) having the production volumes of greater than four billion pounds in 1990, according to the U.S. International Trade Commission Report, are listed below. Note that the production of these 13 chemicals alone totaled more than 100 billion pounds.

Ethylene dichloride	13.8 billion lb/yr
Ethylbenzene	8.4
Styrene	8.0
Urea	12.2
Vinyl chloride	10.6
Methanol (synthetic)	8.3
Formaldehyde (37%)	6.7
p-Xylene	5.2
DMT/Terephthalic acid	7.8
Ethylene glycol	5.1
Ethylene oxide	5.3
MTBE	8.9
Cumene	4.3

*Dow Chemical Co., Midland, MI.
**Union Carbide, So. Charleston, W.V.

Fig. 22.1. Total production of synthetic organic chemicals, 1953–90 (*U.S. International Trade Commission, Dec. 1991.*)

Almost three thousand organic chemical products are currently derived from petrochemical sources. The commercial syntheses for all these products, however, can be traced back to one of six logical starting points. Consequently, this chapter has been subdivided according to the six major raw material sources: chemicals derived from methane, those from ethylene, those from propylene, those from "C_4's", those from higher aliphatics, and those from the aromatics.

A number of general references at the end of the chapter may be consulted for an overview of the subject.

CHEMICALS DERIVED FROM METHANE

Methane is readily available as the major component of natural gas, and its primary use is as fuel. It is also found in coal mines and as a product of anaerobic biological decay of organic materials in swamps and landfills. It is possible to make many chemicals from methane in a laboratory. However, methane is relatively inert chemically and is truly useful as a raw material for only a few commercial chemicals. These conversions generally require high temperatures and pressures or very aggressive chemicals such as chlorine, and usually are operated on a very large scale. Here, only those materials that are currently made from methane in commercial quantities are considered. The most important of these are shown in Fig. 22.2.

Synthesis Gas

The most important route for the conversion of methane to petrochemicals is via either hydrogen or a mixture of hydrogen and carbon monoxide. The latter material is known as synthesis gas. The manufacture of carbon monoxide–hydrogen mixtures from coal was first established industrially by the well known water–gas reaction:

$$C + H_2O \rightarrow CO + H_2$$

Two important methods presently are used to produce the gas mixture from methane. The first is the methane–steam reaction, where methane and steam at about 900°C are passed through a tubular reactor packed with a promoted iron oxide catalyst. Two reactions

Fig. 22.2. Some important synthetic chemicals derived from methane.

are possible, depending on the conditions:

$$CH_4 + H_2O \rightarrow CO + 3H_2$$

$$CO + H_2O \rightarrow CO_2 + H_2$$

The second commercial method involves the partial combustion of methane to provide the heat and steam needed for the conversion. Thus the reaction can be considered to take place in at least two steps, the combustion step:

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$$

followed by the reaction step:

$$CH_4 + CO_2 \rightarrow 2CO + 2H_2$$

$$CH_4 + H_2O \rightarrow CO + 3H_2$$

The process usually is run with nickel catalysts in the temperature range of 800 to 1000°C. Steam reforming usually is used on the lighter feedstocks, and partial oxidation is used for the heavier fraction.

Synthesis gas is the starting material for the manufacture of ammonia and its derivatives and also for methanol, as well as for other oxo-synthesis processes. It also is a source of carbon monoxide in the manufacture of such chemicals as acetic acid and a source of hydrogen for various petroleum refining processes. However, shortages and corresponding high prices of natural gas and naphtha have generated interest in other synthesis gas feedstocks such as coal and residual oil.

Ammonia. Although ammonia is not an organic chemical, it is one of the largest-volume synthetic petrochemicals. From it many reactive organic chemicals are derived such as urea, acrylonitrile, caprolactam, amines, and isocyanates. Almost all of the 30 or more billion lb of ammonia produced in 1988 in the United States was based on hydrogen from petroleum and natural gas. Detailed descriptions of ammonia processes are found in Chapter 28.

Methanol. Methanol was once known as wood alcohol because it was a product of the destructive distillation of wood. Before 1926,

all American methanol was produced in that way. That year, however, marked the first appearance in the world of German synthetic methanol. Today, almost all of the approximately 8.3 billion lb/year of methanol made in the United States comes from large-scale, integrated plants for conversion of natural gas to synthesis gas to methanol (Fig. 22.3).

In principle, methanol, because it is derived from synthesis gas, can be made not only from convenient natural gas but from any source of reduced organic carbon such as coal, wood, or cellulosic agricultural waste. It then can be used as a readily stored fuel or shipped for use as fuel or raw material elsewhere. As petroleum and natural gas become more difficult to recover, alternative carbon sources such as these will be used more. An example involving methanol is a coal-based acetic anhydride facility started in the 1980s in the United States. Another example is the use of some of the natural gas formerly wasted during recovery of Middle Eastern oil to make methanol and other chemicals. Direct use of methanol as a motor vehicle fuel is being studied, but it is not known when or even if such use will be significant in terms of methanol usage or decreased petroleum usage.

Methanol synthesis resembles that of ammonia in that high temperatures and pressures are used to obtain high conversions and rates. Improvements in catalysts allow efficient operation at temperatures and pressures much lower than those of the initial commercial processes. Today, "low-pressure" Cu–Zn–aluminum oxide catalysts are operated at about 1500 psi and 250°C. These catalysts must be protected from trace impurities that the older "high-pressure" (5000 psi and 350°C) and "medium-pressure" (3000 psi and 250°C)

Fig. 22.3. Manufacture of chemical or fuel-grade methanol from hydrocarbon feedstocks using ICI low-pressure process. (Hydrocarbon Processing, *p. 144, Nov. 1985. Copyright 1985 by Gulf Publishing Company and used by permission of the copyright owner.*)

catalysts tolerate better. Synthesis gas production technology has coevolved so that it is possible to maintain the required low levels of these trace impurities.

Methanol is used as a solvent, an antifreeze, a refrigerant, and a chemical intermediate. The greatest chemical uses for methanol as of 1990 were: formaldehyde—27 percent; MTBE—26 percent; and acetic acid—11 percent. Other chemicals derived from methanol include methyl halides, methyl methacrylate, and methylamines. Additional methanol capacity will be required to meet the expected growth of MTBE consumption.

Formaldehyde. Formaldehyde may be made from methanol either by catalytic vapor-phase oxidation:

$$CH_3OH + \tfrac{1}{2}O_2\,(\text{air}) \rightarrow CH_2O + H_2O$$

or by a combination oxidation–dehydrogenation process:

$$CH_3OH \rightarrow CH_2O + H_2$$

It also can be produced directly from natural gas, methane, and other aliphatic hydrocarbons, but this process yields mixtures of various oxygenated materials.

Because both gaseous and liquid formaldehyde readily polymerize at room temperature, formaldehyde is not available in pure form. It is sold instead as a 37 percent solution in water, or in the polymeric form as paraformaldehyde $[HO(CH_2O)_nH]$, where n is between 8 and 50, or as trioxane $(CH_2O)_3$. The greatest end use for formaldehyde is in the field of synthetic resins, either as a homopolymer or as a coplymer with phenol, urea, or melamine. It also is reacted with acetaldehyde to produce pentaerythritol $[C(CH_2OH)_4]$, which finds use in polyester resins. Two smaller-volume uses are in urea–formaldehyde fertilizers and in hexamethylenetetramine, the latter being formed by condensation with ammonia.

U.S. production of formaldehyde in both 1988 and 1989 was approximately 6500 million lb, amounting to about 70 percent of capacity. Usage is expected to grow at about 3 percent/year through 1994.

Methyl Methyacrylate. Methyl methacrylate is formed in a three-step process from methanol, acetone, and HCN:

$$CH_3\overset{\overset{\displaystyle O}{\|}}{C}CH_3 \xrightarrow{HCN} CH_3\underset{\underset{\displaystyle OH}{|}}{\overset{\overset{\displaystyle CH_3}{|}}{C}}CH \xrightarrow{H_2SO_4}$$

$$CH_2{=}\underset{\underset{\displaystyle CH_3}{|}}{C}CONH_2\cdot H_2SO_4 \xrightarrow{CH_3OH}$$

$$CH_2{=}\underset{\underset{\displaystyle CH_3}{|}}{C}COOCH_3 + NH_4HSO_4$$

Although this is the major process in operation, there also is at least one commercially operated process for converting isobutylene and/or *tert*-butanol to methacrylic acid followed by esterification with methanol. A new process based on acetone cyanohydrin also has been reported. It avoids corrosive sulfuric acid and by-product salts while coproducing formamide, which can be converted to HCN for recycle to the process.

U.S. production of methyl methacrylate in 1988 totaled a billion pounds, which is about 90 percent of capacity. Its uses are almost exclusively based on polymerization to poly(methyl methacrylate), which, because of its physical strength, weathering resistance, optical clarity, and high refractive index, has major uses in cast and extruded sheet (25%), molding powders and resins (25%), and surface coatings (20%).

Acetic Acid. Acetic acid production was reported to be 6.57 billion lb in 1988. Of this, U.S. capacity was 3.41, Western European 2.29, and Japanese 0.85 billion lb. Projected worldwide growth from 1988 to 1995 is 2.5 percent/year, with Japanese growth the greatest and European the least.

Acetic acid is produced by methanol carbonylation, which has become the dominant process, as well as by acetaldehyde oxidation and light hydrocarbon oxidation. When methanol carbonylation was first practiced in

the United States and West Germany, a cobalt iodide catalyst was used, and the process required up to 10,000 psia pressure. The technological breakthrough that allowed methanol carbonylation to become the leading acetic acid process was the discovery of rhodium–iodine catalysts, which can be operated at moderate pressure (500 psia) and at a methanol selectivity of 99 percent to acetic acid. Process optimization has minimized the loss of the very costly rhodium catalyst, and has limited interest in developing catalysts based on less expensive metals. Figure 22.4 is a schematic of the process, and Fig. 22.5 shows the catalytic chemistry for acetic acid synthesis by methanol carbonylation.

In the United States, 3.41 billion lb of acetic acid were consumed in 1988. The applications, in decreasing order were: vinyl acetate (which alone accounts for more than 50% of U.S. acetic acid consumption), dimethyl terephthalate/terephthalic acid, acetate esters, cellulose acetate, other acetic anhydride uses, textiles, monochloroacetic acid, and several smaller uses. Growth projections are for an overall 2.7 percent annual growth from 1988 to 1995, with dimethyl terephthalate and terephthalic acid leading the growth and with an actual decline for monochloroacetic acid.

Methyl t-Butyl Ether (MTBE). In 1980, MTBE was the fastest-growing derivative of methanol. In 1990, it was the fastest-growing chemical in the world.[14] This is a result of its only significant use, which is as an antiknock agent replacing lead in gasoline. The growth in U.S. capacity from 1980 through 1985 was 40 percent/year. In 1985, the EPA enacted rules requiring staged decreases in the lead content in motor fuels. The worldwide growth in capacity from 1987 through 1990 was nearly double that of the early eighties, with almost another doubling predicted for 1990 through 1994.

MTBE is made by reacting methanol with the isobutylene contained in mixed-C_4 refinery streams. This is possible because butanes, the other butenes, and butadiene are inert under the mild conditions used. The process is catalyzed by acidic ion exchange resins.

Oxo Chemicals. The so-called oxo process combines carbon monoxide and hydrogen with olefins to make saturated aldehydes having one more carbon atom than the olefins have. The earliest such reaction studied used ethylene to produce both an aldehyde and a ketone. Thus, the reaction was named "oxo" after the German *oxierung*, meaning "ketonization."

The older, high-pressure process, which used cobalt carbonyls as catalyst, has largely been replaced by the low-pressure oxo process based on rhodium complex catalysts. The low-pressure process is operated at about 100°C and 200 psig. This is shown schematically in Fig. 22.6.

Oxo chemicals include butyraldehyde (*normal-* and *iso-*) and the corresponding alcohols, 2-ethylhexanol (from *n*-butyraldehyde), propionaldehyde and *n*-propyl alcohol, and lesser amounts of higher aldehydes and alcohols derived from C_5 through C_{17} olefins. The total volume of products derived from oxo chemistry exceeds a billion pounds a year. Volumes and applications are given later in this chapter for the most important products.

Chloromethanes

The four chlorinated methanes are methyl chloride (CH_3Cl), methylene chloride (CH_2Cl_2), chloroform ($CHCl_3$), and carbon tetrachloride (CCl_4). They are all fairly large-volume products, with their recent U.S. annual production levels being about 520, 500, 450, and 750 million lb respectively.

Methyl chloride currently is produced by two methods: by the reaction of hydrogen chloride and methanol and by the chlorination of methane. Methanol hydrochlorination has become increasingly important, whereas methane chlorination has declined as the route to produce only methyl chloride. The hydrochlorination process has the advantage that it utilizes, instead of generating, hydrogen chloride, a product whose disposal has become increasingly difficult. The reaction of methanol with excess hydrogen chloride is carried out in the gas phase at temperatures

Fig. 22.4. Acetic acid via methanol carbonylation—Monsanto process. (Chem Systems Report No. 88-7. *Copyright Chem Systems Inc., Tarrytown, NY, and used by permission of copyright owner.*)

Fig. 22.5. Cycles for catalytic action of rhodium and iodine in acetic acid synthesis. (Chem Systems Report No. 88-7. *Copyright Chem Systems, Inc. and used by permission of the copyright owner.*)

Fig. 22.6. Butyraldehyde process using triphenylphosphine-modified rhodium carbonyl catalyst. (Encyclopedia of Chemical Technology, *Kirk and Othmer, 3rd ed., Vol. 16, p. 648, 1980. Copyright by John Wiley & Sons, Inc. and reproduced by permission of the copyright owner.*)

of 280 to 350°C:

$$CH_3OH + HCl \rightarrow CH_3Cl + H_2O$$

This route also generates only a single product, rather than the mixture of products obtained from methane chlorination.

Methylene chloride and chloroform can be made, along with the other products, by the direct chlorination of methane. It is much more common, however, to produce them by the chlorination of methyl chloride. This can be done either thermally (350–450°C) or photochemically. The HCl by-product can be recycled back into a hydrochlorination process for production of the methylene chloride starting material.

Until the late 1950s, most carbon tetrachloride was made by the reaction of chlorine and carbon disulfide. At that time commercial processes based on the high-temperature chlorination of methane or the cleavage of C_2–C_3 hydrocarbons and their chlorinated derivatives were introduced. In these processes (known as perchlorinations or chlorinolysis)

the potential starting materials include ethylene, propane, propylene, dichloroethane, dichloropropane, and by-products from other processes such as methane chlorination and vinyl chloride production. A more detailed description of this type of process is included in the section covering vinyl chloride production.

The production of chlorinated methanes peaked in the 1970s, and their future likely will be dominated by environmental concerns about them and some of their major products. The principal use for methyl chloride, other than as a starting material for the other products, is in the manufacture of chlorosilanes for the silicone industry. Other smaller uses are for methylcellulose and tetramethyllead and in miscellaneous methylating reactions. Methylene chloride is used primarily as a solvent for degreasing and paint removal, and it also is used in aerosols and foam-blowing agents. All of these applications are under study, however, because of air pollution problems. Almost all of the chloroform is reacted with hydrogen fluoride to produce chlorodifluoromethane (Fluorocarbon 22). This material is used as a refrigerant and as an aerosol propellant, as well as to synthesize the monomer tetrafluoroethylene. The volume of carbon tetrachloride used has decreased steadily, and its future will be determined mainly by the fate of chlorofluorocarbons as aerosol propellents. About two-thirds of all carbon tetrachloride goes to produce dichlorodifluoromethane (Fluorocarbon 12), and most of the remainder ends up as trichlorofluoromethane (Fluorocarbon 11).

Acetylene

In the early days of the chemical industry, acetylene was a key starting material for many important products. Initially it was obtained for chemical purposes by reaction of calcium carbide with water; but that practice has given way to acetylene recovery from hydrocarbon cracking, so that now 90 percent of acetylene used in chemical manufacturing is made in this way. Owing to difficulty in its safe collection and transport, it almost always is used where it is prepared.

Acetylene still is a preferred raw material for some products, but it has been largely supplanted by ethylene for many others. Chemicals once produced from acetylene by processes now considered outdated include: vinyl chloride, vinyl acetate, acetaldehyde, acrylonitrile, neoprene, and chlorinated solvents. In 1990, an improved, carbide-based acetylene process was being implemented to supply raw material for chemical processes.[15] This may make acetylene more competitive with ethylene as a raw material.

1,4-Butanediol. In 1988, 400 million lb of 1,4-butanediol were made, of which 120 million lb were sold. This is up from 300 million lb reported in 1978. Of the chemical uses of acetylene, 1,4-butanediol is the largest consumer other than VCM (Vinyl Chloride Monomer). The use of acetylene to make VCM is now considered outdated, and it is a minor process compared to the production of VCM from ethylene.

The Reppe process is used to make 1,4-butanediol from acetylene. In this process, acetylene and formaldehyde are reacted in the presence of a copper–bismuth catalyst. The resulting intermediate, 2-butyne-1,4-diol is hydrogenated over Raney nickel:

$$\underset{\text{acetylene}}{HC{\equiv}CH} + \underset{\text{formaldehyde}}{2HCHO} \longrightarrow \underset{\text{2-butyne-1,4-diol}}{HOCH_2C{\equiv}CCH_2OH}$$

$$\underset{\text{2-butyne-1,4-diol}}{HOCH_2C{\equiv}CCH_2OH} + 2H_2 \longrightarrow \underset{\text{1,4-butanediol}}{HO(CH_2)_4OH}$$

Applications of 1,4-butanediol include the manufacture of THF (tetrahydrofuran), used as a solvent, and of poly(butylene terephthalate), used in thermoplastic resins and growing at 5 percent annually.

Hydrogen Cyanide

Most of the hydrogen cyanide used today is prepared, as illustrated in Fig. 22.7, by ammoxidation of methane over a platinum catalyst. Absorption in water and distillation give pure hydrogen cyanide. Although all new hydrogen cyanide capacity is based on this technology, it also can be made from coke-oven gas, from acidification of inorganic cyanides, and by dehydration of formamide. A significant amount of material is obtained as a by-product of ammoxidation of propylene in acrylonitrile manufacture. This amounts to about 25 percent of the demand. Because hydrogen cyanide is very toxic, much of it is used by the producers, to minimize the potential for human exposure during shipping.

The U.S. capacity in 1990 was 1500 million lb. Uses for hydrogen cyanide include: adiponitrile (for nylon 6/6), 43 percent; methyl methacrylate, 33 percent; and various others including chelating agents, contributing to the 1990 U.S. demand of about 1200 million lb.

Carbon Disulfide

Carbon disulfide is made by the catalytic reaction of methane with sulfur vapor. This can be thought of as analogous to oxidation of methane with oxygen, giving carbon dioxide. Carbon disulfide is used as raw material for making carbon tetrachloride and in processes for making rayon, cellophane, and other regenerated cellulosics. Regulatory pressures against chlorofluorocarbons, for which carbon tetrachloride is a raw material, and a decline in the production of rayon and regenerated cellulosics have limited the growth of carbon disulfide uses.

CHEMICALS DERIVED FROM ETHYLENE

Ethylene surpasses all other organic petrochemicals in production and in the amount sold. It is used as raw material for a greater number of commercial synthetic organic chemical products than is any other single chemical. Figure 22.8 shows the more important derivatives of ethylene. Ethylene consumption has grown explosively since 1940 when 300 million lb were used, mostly for making ethanol and ethylene oxide. During World War II, styrene use grew markedly, and polyethylene was developed as insulation for the then-new radar electronics. These materials later found a multitude of applications, which were responsible in large part for ethylene consumption reaching nearly 5 billion lb in 1960. Strong growth in ethylene dichloride and ethylene oxide contributed to over 18 billion lb of ethylene consumption in 1970. Continued growth raised this figure to 27 billion lb in 1978 and over 37 billion lb in 1988, which was more than one hundred times the 1940 volume. The overall growth for 48 years averaged a remarkable 10 percent/year.

Fig. 22.7. The hydrogen cyanide process. (*Reprinted from* Ind. Eng. Chem., ***51**, no. 10, 1235, 1959; copyright 1959 by the American Chemical Society and reprinted by permission of the copyright owner.*)

Fig. 22.8. Important derivatives of ethylene.

Ethylene is manufactured by cracking hydrocarbons. A discussion of hydrocarbon cracking as a route to ethylene is found in Chapter 15.

The major U.S. consumers of ethylene are shown in Fig. 22.9.

Polyethylene

The largest consumers of ethylene are the various types of polyethylene: LDPE (Low Density Polyethylene), HDPE (High Density Polyethylene), and LLDPE (Linear Low Density Polyethylene). Chapter 19 gives detailed discussions of preparation of the various types of polyethylene.

LDPE is produced by high-pressure, high-temperature radical polymerization of pure ethylene. When improved properties are required, copolymers with one or more other vinyl monomers such as ethyl acrylate, vinyl acetate, or acrylic acid are used. LDPE has a relatively branched molecular structure, and the branches are relatively long. It is used for a multitude of purposes because of its properties and economics. Some important uses include: films for packaging of food and other merchandise; shipping trays and pallets; lightweight, flexible, water- and chemical-resistant containers or barriers; and temporary coverings as in construction and agriculture.

HDPE is produced by a low-pressure, fluid-bed polymerization process in which highly pure gaseous ethylene is converted by proprietary catalysts to solid polyethylene particles. It has a very linear molecular structure. HDPE is stronger, tougher, and more rigid than LDPE, so it is used where such properties are advantageous. The major uses for HDPE are in blow-molded bottles, cans, and tanks for products such as milk, bleach, detergent, and fuel and in grocery sacks and other paper-replacement markets.

LLDPE is made by a fluid-bed, catalytic process very similar to that for HDPE, but it is a softer polyethylene than HDPE with

TOTAL: 1988: 37,055 Million Pounds
1989: 35,270 Million Pounds
1990: 37,140 Million Pounds

CHEM SYSTEMS

Fig. 22.9. U.S. ethylene demand, 1990 (*Courtesy Harold A. Witcoff and Chem Systems, Inc.*)

properties similar to those of LDPE. Its properties are achieved by inclusion of comonomers such as butene or hexene. A relatively disordered crystalline state is obtained by introducing many short branches into an otherwise highly linear molecule. Thus, the less expensive equipment of the HDPE process can be used to make a product having the greater flexibility and impact strength characteristic of LDPE.

In 1988 U.S. demand for polyethylene was nearly 19 billion lb. The low-density products (both LDPE and LLDPE) amounted to 10.6 billion lb, and the high-density polyethylene accounted for the other 8.3 billion lb. Together, these products consumed slightly more than half of the 37 billion lb of ethylene produced in the United States.

Ethylene Oxide

Ethylene oxide was discovered in 1859 by Wurtz. He stated that ethylene oxide could not be made by direct oxidation of ethylene, and it was nearly 80 years before this was disproved. Wurtz made ethylene oxide by the method known today as the chlorohydrin process, in which ethylene is reacted in turn with hypochlorous acid and base. This process was commercialized during World War I in Germany, and until 1985 was still used commercially in the United States.

$$CH_2{=}CH_2 + HOCl \longrightarrow CH_2ClCH_2OH$$

$$2CH_2ClCH_2OH + Ca(OH)_2 \longrightarrow CaCl_2 + 2H_2O + 2\,\underset{\diagdown\,O\,\diagup}{CH_2{-}CH_2}$$

Since 1985, processes for the direct oxidation of ethylene using either air or oxygen and a silver catalyst have been the only remaining commercial processes for ethylene oxide production in the United States. Figure 22.10 illustrates an air-based process, and Fig. 22.11 an oxygen-based process. In this very exothermic conversion, oxygen and ethylene combine on a silver catalytic surface to make ethylene oxide. Oxygen and ethylene concentrations are controlled at low levels to avoid creating explosive mixtures.

The competing reactions of total combustion to carbon dioxide and isomerization must be avoided. Ethylene oxide plants in which air

Fig. 22.10. Air-based direct oxidation process for ethylene oxide. (Encyclopedia of Chemical Technology, *Kirk and Othmer, 3rd ed., Vol. 9, p. 443, 1980. Copyright by John Wiley & Sons, Inc. and reproduced by permission of the copyright owner.*)

Fig. 22.11. Oxygen-based direct oxidation process for ethylene oxide. (Encyclopedia of Chemical Technology, *Kirk and Othmer, 3rd ed., Vol. 9, p. 446, 1980. Copyright by John Wiley & Sons, Inc. and reproduced by permission of the copyright owner.*)

is used as source of oxygen require additional investment for purge reactors and associated absorbers. This investment is offset by the need, in the oxygen-based process, for an oxygen production plant and a carbon dioxide removal system. In general, the oxygen-based process is thought to be more economical for smaller plants, and the air-based process is preferred for larger plants, with the breakpoint being between about 100 million and 150 million lb/year.

In 1987, U.S. ethylene oxide capacity was 6.3 billion lb with production of 5.8 billion lb. Major uses in that year were: ethylene glycol—59 percent; nonionic surfactants—14 percent; ethanolamines—8 percent; glycol ethers—6 percent; diethylene glycol—6 percent; and triethylene glycol—2 percent. The remaining 5 percent of ethylene oxide consumption included PEGs (poly-(ethylene glycol)), urethane polyols, and exports. In the following sections, several of these derivatives are discussed in more detail.

Ethylene Glycols. Monoethylene glycol or ethylene glycol is the major derivative of ethylene oxide. Ethylene glycol was initially made commercially by hydrolysis of ethylene chlorohydrin. Today, hydrolysis of ethylene oxide is the preferred route.

$$\overset{O}{\diagup\!\diagdown}\!CH_2CH_2 + H_2O \longrightarrow HOCH_2CH_2OH$$

Reaction of ethylene oxide with water is accomplished in a scheme such as that shown in Fig. 22.12. A large molar excess of water is supplied to the reactor to favor ethylene glycol formation over that of diethylene glycol and triethylene glycol, as the glycols are more reactive with ethylene oxide on a molar basis than is water. A multiple-effect distillation scheme is used to recycle the unreacted excess water. The major product, ethylene glycol, is recovered between light impurities and the heavier glycol by-products.

As a result of the greater reactivity of ethylene glycol with ethylene oxide than that of water with ethylene oxide, it is sometimes difficult to keep production of diethylene glycol from exceeding demand. In this regard, there is technology based on the intermediacy of ethylene carbonate, which allows the manufacturer to limit production of diethylene glycol to that needed because, in contrast to direct ethylene oxide hydrolysis, hydrolysis of ethylene carbonate to ethylene glycol does not produce diethylene glycol.

Ethylene glycol uses are dominated by antifreeze and polyester fiber, at about 43 percent and 26 percent, respectively. Poly-(ethylene terephthalate) or PET resin uses total about an additional 19 percent, with PET bottles being 7 percent. This last use is growing much more rapidly than that of ethylene glycol overall.

Diethylene glycol usage is about 500 million lb/year in the United States, roughly 10

Fig. 22.12. Production of ethylene glycols from ethylene oxide. (1) Feed tank containing ethylene oxide and water; (2) hydration reaction; (3, 4, 5) multiple stage evaporator; (6) light ends column; (7, 8) vacuum distillation towers. (Hydrocarbon Processing, Petrochemicals Handbook, *Nov. 1989, p. 102. Copyright Gulf Publishing Company and reproduced by permission of the copyright owner.*)

percent of that of ethylene glycol. Major uses are 34 percent in unsaturated polyester resins and polyurethane resins and 17 percent in antifreeze blending. Other uses include 13 percent as raw material for triethylene glycol and 8 percent as raw material for morpholine manufacture. It also is used for dehydration of natural gas and in textile conditioning.

Triethylene glycol consumption, at about 120 million lb/year year in the United States, is roughly 2 percent of that of ethylene glycol. The major use in natural gas drying depends on triethylene glycol's low volatility and strong affinity for water. Lesser amounts are used as intermediate for vinyl plasticizers, polyester resins, and polyols. Additional direct applications include solvent and humectant uses.

Polyethylene glycols are produced by base-catalyzed addition of ethylene oxide to a low molecular weight glycol such as diethylene glycol. These glycols are higher-molecular-weight analogs of mono-, di-, and triethylene glycol but differ from the latter compounds in that they are not pure substances but rather consist of distributions of low molecular weight polymers. With average molecular weights beginning at about 200 (that of tetraethylene glycol) and going up to about 1000, these materials are liquids at ambient temperature. They are used as plasticizer intermediates, dispersant media, lubricants, and humectants. Above an average molecular weight of 1000, the polyglycols become waxy solids and find use in ointments, cosmetics, and lubricants taking advantage of their oil and water compatibility and low toxicity. At very high molecular weights, homopolymers of ethylene oxide are used for thickening, for water-soluble films, and for reducing friction in, for example, water delivery in fire hoses. Their value in this last application is that a given size hose can be made to deliver a greater flow of water.

Surfactants. Ethylene oxide–containing surfactants are generally of nonionic or anionic classes. The nonionic materials are made by base-catalyzed addition of ethylene oxide to either fatty alcohols or alkylphenols. Sulfation can be used to convert these compounds to the sulfated anionic surfactants. The products contain from a few to many ethylene oxide molecules per alcohol. The chain of poly-(ethylene oxide) in a nonionic product acts as the hydrophile, and the alkyl or alkaryl residue is the hydrophobe. A sulfate salt group adds to the hydrophilicity of an anionic surfactant.

Surfactants based on aliphatic alcohols are used as cleaners in both domestic and industrial applications. They provide excellent properties such as wetting, dispersion, and emulsification. The ethoxylates derived from alkylphenols are chemically stable and highly versatile, finding more use in industrial practice than in domestic applications. They are used both as processing aids and as components in various products. Their applications include metal cleaning, hospital cleaners and disinfectants, agricultural chemical formulation surfactants, insecticides and herbicides, oil-well drilling fluids, and many others.

Examples of volumes used can be seen by considering two major classes of materials. About 400 million lb of nonylphenol ethoxylates and about 640 million lb of mixed linear alcohol ethoxylates were produced in 1988 in the United States.

Ethanolamines. Ethanolamines are manufactured by reacting ethylene oxide and ammonia. The relative amounts of the three amines will depend primarily on the ammonia-to-oxide feed ratio.

$$NH_3 \xrightarrow{C_2H_4O} (HOC_2H_4)NH_2 \xrightarrow{C_2H_4O}$$

$$(HOC_2H_4)_2NH \xrightarrow{C_2H_4O} (HOC_2H_4)_3N$$

The three products are separated by distillation. Over the years, the relative demand for the three products has varied greatly. Thus, operational flexibility must be maintained.

The ethanolamines are water-miscible bases from whose properties stem their major uses as neutralizers in aqueous formulations such as metalworking fluids. Monoethanolamine is used in detergents, in "sweetening" (removing

carbon dioxide and hydrogen sulfide from) natural gas, for removing carbon dioxide from ammonia during its manufacture, and as a raw material for producing ethyleneamines by reductive amination. Diethanolamine finds use in detergents and as an absorbent for acidic components of gases, as well as its major use as a raw material for surfactant diethanolamides of fatty acids. Triethanolamine's main end uses are in cosmetics and textile processing.

Production capacity (U.S.) for ethanolamines in 1989 was about 900 million lb, and demand was 650 million lb. Overall in 1989, 23 percent of ethanolamines went to surfactants, 15 percent to natural gas conditioning, and percent to 7 metalworking. Other uses in the United States accounted for 21 percent, and exports for 34 percent.

Glycol Ethers. In the same way that water reacts with one or more molecules of ethylene oxide, alcohols react to give monoethers of ethylene glycol, producing monoethers of diethylene glycol, triethylene glycol, and so on, as by-products.

Since their commercial introduction in 1926, glycol ethers have become valuable as industrial solvents and chemical intermediates. Because glycol monoethers contain a $—OCH_2CH_2OH$ group, they resemble a combination of ether and ethyl alcohol in solvent properties. The most common alcohols used are methanol, ethanol, and butanol. Principal uses for the glycol ethers are as solvents for paints and lacquers, as intermediates in the production of plasticizers, and as ingredients in brake fluid formulations. Condensation of the monoethers produces glycol diethers, which are also useful as solvents.

$$2ROCH_2CH_2OH \xrightarrow{H_2SO_4} ROC_2H_4OC_2H_4OR + H_2O$$

Solvent characteristics of glycol ethers are enhanced by esterifying with acetic acid. The resulting acetate esters are used extensively in coating formulations, especially those formulations in which their high solvent power allows a decreased total solvent usage in compliance with VOC (Volatile Organic Compound) emission standards.

In recent years, health concerns have caused substitution of propylene glycol–based products for the ethylene glycol–based products to avoid human exposure.[16]

Other Uses of Ethylene Oxide. About 2 percent (120 million lb, U.S., 1987) of ethylene oxide is consumed in miscellaneous applications, such as its use as a chemical intermediate for, for example, choline and ethylene chlorohyrin, hydroxyethyl starch, and hydroxyethyl cellulose and its direct use as a fumigant/sterilant. As environmental and health considerations make it more difficult to transport toxic chemicals, ethylene oxide uses tend more toward those in which the manufacturer uses the material at the site of manufacture and away from uses of the pure material, such as in fumigation.

Chlorinated Ethanes and Ethylenes

A number of important large-volume petrochemicals are obtained through the chlorination of ethane and ethylene. The largest-volume chlorinated derivative, 1,2-dichloroethane, has a production volume of more than 13 billion lb/year. Because of their unique solvent and chemical intermediate properties, the market for chlorinated ethanes and ethylenes grew steadily until it reached a peak in the late 1970s. Owing to environmental problems, particularly in the solvent area, the demand for some of the end uses has been declining steadily since then. Figure 22.13

Fig. 22.13. Manufacturing processes for C_1 and C_2 chlorohydrocarbons. (Encyclopedia of Chemical Technology, *3rd ed., Vol. 5, p. 671. Copyright by John Wiley & Sons, Inc. and reproduced by permission of the copyright owner.*)

shows the possible production routes to the major chlorinated derivatives.

Chlorinated Ethanes. Of the nine possible chlorinated derivatives of ethane, only three are of commercial importance: ethyl chloride, 1,2-dichloroethane (ethylene dichloride), and 1,1,1-trichloroethane (methyl chloroform). The other compounds have no important end uses and are produced either as intermediates or as unwanted by-products. They normally are converted to useful materials by a cracking process (for trichloroethylene) or by perchlorination (for carbon tetrachloride and tetrachloroethylene).

Ethyl Chloride. Most of the ethyl chloride is made by the exothermic hydrochlorination of ethylene, in either the liquid or the vapor phase:

$$CH_2{=}CH_2 + HCl \xrightarrow{AlCl_3} CH_3CH_2Cl$$

A much smaller amount is produced by the thermal chlorination of ethane. This direct chlorination may be run in conjunction with another process, such as oxychlorination, that can use the by-product HCl as feed.

Ethyl chloride rose to commercial importance because of the automotive industry. It was the starting material for tetraethyllead, at one time the most commonly used octane booster. Demand has been cut drastically because of the conversion from leaded to unleaded gasoline for environmental reasons. Production of ethyl chloride in 1988 was estimated to be about 180 million lb, less than a third of what it had been ten years previously. Smaller uses for ethyl chloride are in the production of ethyl cellulose, as an ethylating agent, as a blowing agent, and in solvent extraction.

1,2-Dichloroethene (Ethylene Dichloride). Ethylene dichloride (EDC) is one of the truly major petrochemicals. The 1990 Trade Commission Report ranks it in the top five synthetic organic chemicals with production of more than 13 billion lb. Almost 95 percent of this went into the manufacture of vinyl chloride monomer.

Figure 22.14 shows an integrated plant for producing ethylene dichloride and vinyl chloride from ethylene, chlorine, and air. In this process, vinyl chloride (VCM) is produced by the thermal cracking of EDC. The feed EDC may be supplied from two sources. In the first source, ethylene and chlorine are reacted in essentially stoichiometric proportions to produce EDC by direct addition. In the second source, ethylene is reacted with air and HCl by the oxychlorination process. Ideally, both processes are carried out in balance, and the oxychlorination process is used to consume the HCl produced in the cracking and direct chlorination steps. The chemical reactions are as follows:

$$CH_2ClCH_2Cl \rightarrow HCl + CH_2{=}CHCl$$

$$CH_2{=}CH_2 + Cl_2 \rightarrow CH_2ClCH_2Cl$$

$$CH_2{=}CH_2 + 2HCl + \tfrac{1}{2}O_2 \rightarrow CH_2ClCH_2Cl + H_2O$$

Thus the overall reaction for the integrated plant is:

$$4CH_2CH_2 + 2Cl_2 + O_2 \rightarrow 4CH_2{=}CHCl + 2H_2O$$

The direct chlorination of ethylene usually is run in the liquid phase and is catalyzed with ferric chloride. High-purity ethylene normally is used to avoid product purification problems. The cracking (pyrolysis) of EDC to vinyl chloride typically is carried out at temperatures of 430 to 530°C over a catalyst. The hot gases are quenched and distilled to remove HCl and then VCM. The unconverted EDC is returned to the EDC purification train. The oxychlorination step is the heart of the process and has two major variables, the type of reactor and the oxidant. The reactor may be either a fixed bed or a fluidized bed, and the oxidant is either air or oxygen. The temperature is in the range of 225 to 275°C with a copper chloride–impregnated catalyst.

Profitable disposal of the by-product HCl once was the major restriction to the growth of ethylene dichloride as the source of vinyl

Fig. 22.14. Integrated ethylene dichloride/vinyl chloride process. (Hydrocarbon Processing, *p. 174. Nov. 1985. Copyright Gulf Publishing Company and used by permission of the copyright owner.*)

chloride and other chlorinated hydrocarbons. The advent of oxychlorination, which uses the HCl and air to chlorinate, opened the door for the rapid replacement of the acetylene-based routes. The HCl reacts with oxygen in the oxychlorination and generates chlorine *in situ*, which then reacts with the ethylene.

Almost 95 percent of all ethylene dichloride goes to make vinyl chloride. Of that, less than 20 percent actually is isolated as EDC. Smaller uses are as a solvent and as a raw material for other chlorinated hydrocarbons such as trichloroethylene and perchloroethylene. Also a small amount is used to produce ethylene amines.

Vinyl Chloride. Approximately 10.5 billion lb of vinyl chloride (VCM) were produced in the United State in 1988, making it one of the largest-volume petrochemicals. It has been reported that more than 25 percent of the global production of chlorine goes to the manufacture of VCM. Although most of the vinyl chloride comes from ethylene dichloride by the route described previously, it can be obtained from other sources, including its production in the catalytic hydrochlorination of acetylene, and its origin as a by-product in the synthesis of other halogenated hydrocarbons.

More than 95 percent of all VCM is used to produce polyvinyl chloride (PVC), an important polymer for the housing and automotive industries. (A detailed description of PVC is included in Chapter 19.) The rest of the vinyl chloride goes into the production of chlorinated solvents.

1,1,1-Trichloroethane (Methyl Chloroform). Trichloroethane is a major solvent, particularly

for cold and vapor degreasing. Domestic production in 1988 was about 700 million lb, making it the largest-volume chlorinated solvent. It can be produced by three methods: by chlorination of 1,1-dichloroethane, from 1,1,2-trichloroethane via 1,1-dichloroethylene, and by direct chlorination of ethane. In the United States about 70 percent is produced by the first route. In this process the ethylene dichloride feedstock is rearranged to 1,1-dichloroethane via cracking to vinyl chloride, followed by addition of HCl in the presence of a catalyst. For the final step, the dichloroethane is thermally or photochemically chlorinated. The reactions are as follows:

$$CH_2Cl—CH_2Cl \rightarrow CH_2{=}CHCl + HCl$$

$$CH_2{=}CHCl + HCl \rightarrow CH_3—CHCl_2$$

$$CH_3—CHCl_2 + Cl_2 \rightarrow CH_3—CCl_3 + HCl$$

Trichloroethane is used as a solvent in many industrial applications, such as vapor degreasing, cold cleaning, metal cutting fluids, and solvents for adhesives. Other applications are in textile processing, dry cleaning, and the production of printed circuit boards. It has the potential to replace trichloroethylene and perchloroethylene in some solvent applications. Its future growth probably will be adversely affected by stringent regulations restricting losses from vapor degreasing and other units.

Chlorinated Ethylenes. Vinyl chloride (VCM) is by far the largest-volume chlorinated ethylene derivative. The others of commercial interest are tetrachloroethylene (perchloroethylene), 1,2-dichloroethylene, trichloroethylene, and 1,1-dichloroethylene (vinylidene chloride).

Tetrachloroethylene (Perchloroethylene). Perchloroethylene historically has been the dominant solvent in the dry-cleaning industry. Its 1988 domestic sales were about 500 million lb. Although perchloroethylene is a desirable solvent because of its good stability and low flammability, environmental concerns are expected to have a significant impact on its future use.

An oxychlorination/oxyhydrochlorination process for the production of perchloroethylene and trichloroethylene is shown in Fig. 22.15. The process can accept a wide range of low-cost feedstocks, such as ethylene, chlorinated C_2 hydrocarbons, and by-product streams from vinyl chloride, chloromethanes, methyl chloroform, and ethylene dichloride plants. The product ratio of trichloroethylene to perchloroethylene can be adjusted over a wide range.

Other Chlorinated Ethylenes. Trichloroethylene was a major solvent for degreasing in the late 1960s and early 1970s. Since that time, its production has decreased significantly because of environmental pressures on the solvent users and replacement by 1,1,1-trichloroethane. This decline is expected to continue.

Although 1,1-dichloroethylene (vinylidene chloride) is a relatively small-volume product, it provides a way of upgrading the unwanted 1,1,2-trichloroethane by-product from the manufacture of ethylene dichloride and 1,1,1-trichloroethane. Its major use is as an intermediate for polyvinylidene chloride and its copolymers, which are important barrier materials for food packaging. The remainder goes to the manufacture of 1,1,1-trichloroethane.

Ethanol

Ethanol is made by both ethylene hydration and fermentation of starches and sugars. In this section the synthetic route will be discussed. The fermentation route is covered in Chapter 24.

In the World War II era, 72 percent of U.S. ethanol was derived from molasses fermentation. By 1978 the balance was 90 percent from direct catalytic hydration and the rest from fermentation. In 1988 the balance had returned to the dominance of fermentation, with 82 percent of the 7.6 billion lb of U.S. ethanol made in this way. The recent swing toward fermentation is due to the use of 90 percent of the fermentation ethanol as motor fuel, as

Fig. 22.15. Perchloroethylene/trichloroethylene plant. (*Reproduced from* Hydrocarbon Processing, *p. 154, Nov. 1985. Copyright 1985 by Gulf Publishing Co. and used by permission of the copyright owner.*)

Fig. 22.16. Manufacture of ethanol by direct hydration of ethylene. (*Courtesy Gulf Publishing Company.*)

a result of post–oil-embargo U.S. government policy.

Direct hydration of ethylene is by far the major route to synthetic ethanol. It is accomplished under pressure at 250 to 300°C over an acidic catalyst. Figure 22.16 shows such a unit. Ethylene and high temperature steam are mixed and passed over an acidic catalyst, usually phosphoric acid on a support. A modest conversion is achieved even with the severe conditions. Cooling of the exit stream and passage through a separations system give ethylene and water for recycle. Ethanol is made as either a 95 percent azeotrope with water or as an anhydrous material from a drying system.

Synthetic ethanol has the following uses: as a chemical intermediate (for ethyl acetate, ethyl acrylate, glycol ethers, ethylamines, etc.)—30 percent; in toiletries and cosmetics—20 percent; as a coatings solvent—15 percent; as a raw material for vinegar—10 percent; in household cleaners—7 percent; in detergents—5 percent; in pharmaceuticals—5 percent; in printing inks—3 percent; and in miscellaneous uses—5 percent.

Ethylbenzene

Ethylbenzene is used almost exclusively (99%) as a raw material for producing styrene. The remainder is used as solvent and in the manufacture of diethylbenzene. In 1989 demand was about 95 percent of the U.S. capacity of 10 billion lb, and a growth of about 3 percent/year was expected for the next few years.

Over 90 percent of all ethylbenzene is produced by alkylation of benzene with ethylene in the presence of an acidic catalyst such as aluminum chloride (liquid phase process) or an acidic zeolite (vapor phase):

$$C_6H_6 + CH_2{=}CH_2 \rightarrow C_6H_5CH_2CH_3$$

Conversion to styrene is accomplished either by dehydrogenation:

$$C_6H_5CH_2CH_3 \rightarrow C_6H_5CH{=}CH_2 + H_2$$

or by a sequence of oxidation to ethylbenzene hydroperoxide, reduction to methyl phenyl carbinol (by a process that also oxidizes propylene to propylene oxide), and dehydration of the methyl phenyl carbinol to styrene. Figure 22.17 shows an integrated ethylbenzene/styrene plant based on aluminum chloride–catalyzed alkylation of benzene with ethylene followed by dehydrogenation of the resulting ethylbenzene to styrene.

Acetaldehyde, Acetic Acid, Acetic Anhydride, Vinyl Acetate

Acetaldehyde. Acetaldehyde has been made from ethanol by dehydrogenation and by catalytic hydration of acetylene. Today direct oxidation of ethylene in the liquid phase

Fig. 22.17. Integrated plant for manufacture of ethylbenzene and styrene. (*Reproduced from* Hydrocarbon Processing, Petrochemical Handbook, *p. 169, Nov. 1985. Copyright Gulf Publishing Co. and used by permission of the copyright owner.*)

Fig. 22.18. One-stage process for direct oxidation of ethylene to acetaldehyde. (a) Reactor; (b) separating vessel; (c) crude aldehyde tank; (d) scrubber; (e) crude aldehyde tank; (f) cycle-gas compressor; (g) light-ends distillation; (h) condensers; (i) purification column; (l) product cooler; (m) regeneration. (Ullmann's Encyclopedia of Industrial Chemistry, *W. Gerhartz, Ed., 5th ed., Vol. A1, p. 37, 1985, VCH Verlagsgesellschaft, Weinheim, Federal Republic of Germany. By permission.*)

catalyzed by palladium and copper has replaced these earlier methods. Figure 22.18 shows an ethylene-to-acetaldehyde unit based on this last route.

Acetaldehyde once was widely used as raw material for a variety of large-volume chemical products such as acetic acid and butanol. U.S. usage peaked in 1969 at 1.65 billion lb. Today, most of the former uses have been superseded by routes based on C_1 or other chemistry such as methanol carbonylation to acetic acid and butanol from propylene by oxo chemistry. Of the remaining uses, which totaled about 650 million lb in the United States in 1986, acetic acid still is the major consumer at 50 percent. It also is used as a raw material for pyridine and substituted pyridines, peracetic acid, lactic acid, crotonaldehyde, and 1,3-butylene glycol.

Acetic Acid. Acetic acid used to be derived from ethylene with acetaldehyde as an intermediate. The relatively high price of ethylene compared to methanol and carbon monoxide (raw materials for the currently favored process) has caused a shift away from ethylene as a raw material for acetic acid production, as was discussed earlier in the chapter.

Some acetic acid still is derived from ethylene-based acetaldehyde by oxidation with molecular oxygen. This is a free radical process and typically is manganese salt–catalyzed. Peracetic acid, an intermediate whose formation and destruction are controlled by the catalyst, has been used commercially for coproducing epoxides as well as for making acetic acid.

Acetic Anhydride. A total of 2.65 billion lb of acetic anhydride was produced in the United States (1.59), Western Europe (0.73), and Japan (0.33) in 1988. Commercial production of acetic anhydride currently is accomplished through two routes, one involving ketene and the other methyl acetate carbonylation. A former route based on liquid phase oxidation of acetaldehyde is now obsolete.

In the ketene process, acetic acid is thermally dehydrated at 750°C to ketene. The ketene is separated from by-product water and reacted with another mole of acetic acid to produce acetic anhydride. Figure 22.19 is a schematic diagram of this process.

The methyl acetate carbonylation process was successfully started and operated in the early 1980s as part of a coal-to-syngas-to-acetic anhydride complex. This new process introduction resulted in a major improvement in acetic anhydride production economics. In this process, methyl acetate, itself the product of a one-step esterification of acetic acid and methanol, is reacted with carbon monoxide in the presence of a promoted rhodium–iodide catalyst. Figure 22.20 illustrates this process

Fig. 22.19. Acetic anhydride from acetic acid—ketene process. (Chem Systems Report No. 88-7. *Copyright Chem Systems, Inc. and used by permission of the copyright owner.*)

Fig. 22.20. Acetic anhydride by carbonylation of methyl acetate, Halcon/Eastman process. (Chem Systems PERP Report No. 88-7. *Copyright Chem Systems, Inc. and used by permission of the copyright owner.*)

exclusive of esterification to make methyl acetate.

The greatest use of acetic anhydride is in esterifying cellulose to cellulose acetate for application as cigarette filter tow and in textiles. In the United States, acetic anhydride is manufactured by cellulose acetate manufacturers and largely used internally. Other products using acetic anhydride as a raw material are mostly mature with low growth rates; they include triacetin, plastic modifiers and intermediates for pharmaceuticals, herbicides, pesticides, and dyes for polyolefins. Growth in production was projected to be 2.1 percent/year in the United States and less in Western Europe and Japan from 1988 to 1995.

Vinyl Acetate. Vinyl acetate production is the largest consumer of acetic acid worldwide. In the United States, vinyl acetate demand in 1988 was 1.85 billion lb. Growth for the United States in the period 1990 to 1995 was expected to be from 3.0 to 3.5 percent/year and for the Western world, 2.7 percent/year.

Production of vinyl acetate is based mostly on oxidative addition of acetic acid to ethylene in the vapor phase. There is still some vinyl acetate production based on addition of acetic acid to acetylene, but the declining availability of acetylene and its higher cost relative to ethylene have caused this latter route largely to be abandoned. Figure 22.21 illustrates the ethylene-based method.

Vinyl acetate, also known as VAM (for vinyl acetate monomer) finds exclusive use as a monomer or raw material for polymers and copolymers; and latex paints are the largest use for PVA (poly(vinyl acetate)) emulsions. Because latex paints cure without appreciable solvent emissions, regulatory pressures against such emissions favor the use of latex paints over solvent-based coatings. Adhesives are the second largest consumers of PVA emulsions, with a range of applications from packaging and wallboard to consumer "white" glue.

In the second major use of VAM, PVA is converted to PVOH (poly(vinyl alcohol)) by a transesterification reaction with methanol, giving methyl acetate as coproduct. PVOH finds its major end use in textile sizing and adhesives. Further reaction of PVOH with butyraldehyde or formaldehyde gives polyvinyl butyral (PVB) or polyvinyl formal, which together constitute the third largest consumption of VAM. PVB is used almost exclusively in the adhesive laminating inner layer in safety glass. This market has been growing more rapidly than either the PVA or PVOH markets and is expected to continue to do so.

Ethylene Oligomers (alpha olefins) and Linear Primary Alcohols

Linear primary alcohols and alpha olefins in the C-6 to C-18 range have enjoyed remarkable growth in the last three decades. As esters, the C-6 to C-10 alcohols are used for plasticizing poly(vinyl chloride). In the C-12 to C-18 range, the alcohols are used to make readily biodegradable surfactants of various types such as ethoxylates (nonionic), alcohol sulfates, and sulfates of ethoxylates (anionic). Alpha olefins are used as polyethylene comonomer (28%) and as raw materials for detergent alcohols (20%), oxo alcohols (11%), and lubricants and lube oil additives (10%).

Production of linear primary alcohols and production of alpha olefins are accomplished by similar reactions in which ethylene is oligomerized by organometallic catalysts based on aluminum alkyls such as triethylaluminum. The two processes are distinguished by the way in which the growing hydrocarbon chain is removed from the catalyst center. In the case of the alpha olefin products, ethylene growth to a hydrocarbon chain of a few or many two-carbon units is interrupted when the hydrocarbon on aluminum is displaced as an alpha olefin by exchange with fresh ethylene. Thus the aluminum alkyl is regenerated, and chain growth starts again. In production of alpha alcohols, the hydrocarbon group on the aluminum catalyst is cleaved by oxygen at the sensitive carbon–aluminum bond to give aluminum oxide and an alcohol. Figures 22.22 and 22.23 illustrate production schemes for alpha alcohols and alpha olefins, respectively.

Fig. 22.21. Vapor-phase vinyl acetate from ethylene process. (Chem Systems Report No. 88-7. *Copyright Chem Systems, Inc. and used by permission of the copyright owner.*)

Fig. 22.22. Alpha alcohols by Alfol® process. (*Reproduced from* Hydrocarbon Processing, ***50**, no. 11, 126, 1971; copyright 1971 by Gulf Publishing Co.*)

Fig. 22.23. Alpha-olefins process. (*Reproduced from* Hydrocarbon Processing, ***50**, no. 11, 127, 1971; copyright 1971 by Gulf Publishing Co.*)

Demand for alpha olefins in the United States in 1989 totaled nearly 1.5 billion lb, compared with a capacity of about 3 billion lb.

Ethylene–Propylene Elastomers

Ethylene-propylene copolymer and terpolymer rubbers (EPRs) are produced at the rate of about a billion pounds per year worldwide. They ranked behind only styrene-butadiene, carboxylated styrene–butadiene, and polybutadiene rubbers in consumption in 1988. Worldwide capacity, exclusive of formerly Communist producers, has been reported as 1.47 billion lb. Of this, 40 percent is that of five producers in the United States.[17]

EPR is produced by polymerization of a mixture of ethylene and propylene and

Fig. 22.24. Manufacture of ethylene–propylene copolymers (EPM rubber) and ethylene–propylene–diene terpolymers (EPDM rubber). (*Reproduced from* Hydrocarbon Processing, *p. 101, Nov. 1983; copyright 1983 by Gulf Publishing Co.*).

optionally a small amount of a nonconjugated diene such as ethylidene norbornene, norbornene, 1,4-hexadiene, or dicyclopentadiene. Two processes, one a solution and the other a suspension process, are employed. They use organometallic catalysts, the most common being products of combining (organo)-vanadium halides with alkyl aluminum or alkyl aluminum halides. The resulting catalysts are deactivated by water and alcohols. A process schematic is shown in Fig. 22.24.

The comonomer diene confers sulfur vulcanizability on the elastomer. Otherwise, a peroxide cure is required for cross-linking. The polymers are readily oil-extended with 20 to 50 percent oil for many applications. Some uses result from these products' ability to resist oxidation by ozone.

Applications of ethylene–propylene copolymers and terpolymers include: automotive (the major use area), thermoplastic olefin elastomers, single-ply roofing, viscosity index improvers for lube oils, wire and cable insulation, hose, appliance parts, and polymer modification.

Propionaldehyde

Propionaldehyde is produced by the oxo reaction of ethylene with carbon monoxide and hydrogen. *n*-Propyl alcohol is produced by hydrogenation of propionaldehyde, and propionic acid is made by oxidation of propionaldehyde. For a drawing of a process similar to the propionaldehyde process see Fig. 22.6.

n-Propyl alcohol is used as solvent in printing inks and as an intermediate in the preparation of agricultural chemicals. Propionic acid is used as a grain preservative as, for example, in preventing spoilage of wet corn used as animal feed. The use of propionic acid as a grain preservative is an alternative to drying by heating, which consumes fuel, and is considered mostly when fuel is expensive.

Other Ethylene Uses

Some lesser-volume ethylene uses are in:

- *Copolymerization with propylene* to improve impact resistance of grades used in injection molding, wire and cable, and film.
- *Copolymerization in vinyl ester adhesive resins* for special properties.
- *Vinyl ester co- and terpolymers for coating resins* to give special properties for some uses.
- *Vinyl toluene* for use in alkyds (29 million lb of resin in 1988).
- *Aluminum alkyls* used in making organometallic catalysts and as initiators for processes such as ethylene–propylene rubber, polybutadiene, low pressure polyethylene, and ethylene oligomerization to make alpha-olefins and C_6 to C_{18} alcohols.
- *Agriculture*, as a ripening agent for fruits and vegetables.
- *Alkylation of anilines* for chemical intermediates used in pesticides, pharmaceuticals, dyes, and urethane comonomers.
- *Diethyl sulfate* made from sulfuric acid and ethylene and used as an alkylating agent in many applications.

CHEMICALS DERIVED FROM PROPYLENE

Propylene demand for chemical synthesis in 1988 in the United States was 20 billion lb. This demand was exceeded by that of only one other synthetic organic chemical, ethylene. The demand was projected to grow at about 3.5 percent/year through 1995. Major uses of propylene are in polypropylene, acrylonitrile, propylene oxide, and cumene. A breakdown of propylene consumption by product is:

Polypropylene	38%
Acrylonitrile	14
Propylene oxide	11
Cumene	10
Oxo alcohols	8
Isopropyl alcohol	6
Oligomers	5
Acrylic acid	3
Export, other	5

Propylene is produced as a coproduct of ethylene cracking and is a product of petroleum refinery operations.

Polypropylene

More than 30 years after its introduction, polypropylene is the largest chemical consumer of propylene in the United States. It is produced primarily by a bulk or gas phase process, with the older slurry process still used some. Because of its greater stiffness relative to polyethylene, polypropylene is used for more demanding applications. This stiffness is not without a drawback in the form of increased brittleness, which can be moderated through incorporation of ethylene by copolymerization. Its major uses are in fibers and injection molding. Chapter 19 reviews polypropylene in some detail.

Acrylonitrile

In 1960, almost all of the 260 million lb annual production of acrylonitrile was based on acetylene. Ten years later, the volume had risen to 1.1 billion lb, which was based almost entirely on an ammoxidation process with ammonia, propylene, and air as feeds. However, in the latter 1980s the growth rate had slowed considerably.

The air, ammonia, propylene process for acrylonitrile is shown in Fig. 22.25. The main reaction is given below.

Fig. 22.25. Simplified diagram of the Sohio acrylonitrile process. (a) Fluidized-bed reactor; (b) absorber column; (c) extractive distillation column; (d) acetonitrile stripping column; (e) lights fractionation column; (f) product column. (Ullmann's Encyclopedia of Industrial Chemistry, *W. Gerhartz, Ed., 5th ed., Vol. A1, p. 179, 1985. VCH Verlagsgesellschaft, Weinheim, Federal Republic of Germany.*)

In this process the highly exothermic oxidation is performed in a fluidized bed to facilitate heat removal. Note also that acetonitrile and hydrogen cyanide are by-products. In the case of hydrogen cyanide, this source is of major commercial importance.

$$2H_2C{=}CHCH_3 + 2NH_3 + 3O_2 \rightarrow 2H_2C{=}CHCN + 6H_2O$$

Acrylic fibers are by far the major end use for acrylonitrile. They find use primarily in fabrics for clothing, furniture, draperies, and carpets. The second largest consumer of acrylonitrile is ABS (acrylonitrile butadiene styrene) and SAN (styrene acrylonitrile) resins. ABS is useful in industrial and construction applications, and the superior clarity of SAN makes it useful in plastic lenses, windows, and transparent household items.

Propylene Oxide

Propylene oxide was manufactured by the chlorohydrin route first during World War I in Germany by BASF and others. This route (below) involves reaction of propylene with hypochlorous acid followed by treatment of the resulting propylene chlorohydrin with a base such as caustic or lime. The products of the second reaction are propylene oxide and sodium or calcium chloride. Figure 22.26 is a diagram of the chlorohydrin process.

$$H_2C{=}CHCH_3 + HOCl \longrightarrow ClH_2CCHOHCH_3$$

$$ClH_2CCHOHCH_3 + NaOH \longrightarrow H_2C\overset{O}{\diagup\diagdown}CHCH_3 + NaCl$$

More recently, propylene oxide has been made commercially by a process involving catalyzed epoxidation of propylene with *tert*-butyl hydroperoxide or ethylbenzene hydroperoxide.

$$RH + O_2 \rightarrow ROOH$$

$$ROOH + H_2C{=}CHCH_3 \rightarrow ROH + H_2C\overset{O}{\diagup\diagdown}CHCH_3$$

Fig. 22.26. Chlorohydrin process. (Encyclopedia of Chemical Technology, *Kirk and Othmer, 3rd ed., Vol. 19, p. 255, 1980. Copyright by John Wiley & Sons, Inc. and used by permission of the copyright owner.*

These processes convert propylene to its epoxide while reducing the hydroperoxide to the corresponding alcohol (e.g., *tert*-butyl alcohol or phenyl methyl carbinol). Because the processes produce the alcohols in larger amounts than propylene oxide, their success depends upon finding uses for the alcohols. *tert*-Butyl alcohol can be dehydrated to isobutylene and hydrogenated to isobutane for recycle to the propylene oxide process. It also can be converted to MTBE. Phenyl methyl carbinol can be dehydrated to styrene, making this process a more involved conversion of ethyl benzene to styrene than is direct dehydrogenation. Figure 22.27 shows the hydroperoxide process.

Some propylene oxide production is based on peracetic acid epoxidation of propylene, but this is relatively minor compared to the chlorohydrin and hydroperoxide processes.

Major uses for propylene oxide are in the production of urethane polyether foams and propylene glycols. It also is used to make glycol ethers, polyglycols, glycerine, surfactants, and amino propanols.

Propylene Glycols. Propylene oxide is converted to mono- and diglycols by a hydrolysis and refining scheme similar to that for ethylene glycol and diethylene glycol. The propylene glycols are used for many of the same applications as the corresponding products derived from ethylene oxide. Because of their very low order of toxicity, they also can be used for pharmaceutical, cosmetic, and food applications. In 1988, the U.S. consumption of these materials was 840 million lb of mono- and 80 million lb of dipropylene glycol.

Isopropyl Alcohol

Isopropyl alcohol (IPA) has been called the first petrochemical. Both historically and today, it is prepared by sulfuric acid–mediated indirect hydration of propylene (see equations below and Fig. 22.28). Originally it was the source of most of the acetone used in the world. Now, this route must compete with acetone derived from the cumene oxidation process, in which cumene is converted to equimolar amounts of phenol and acetone. Between 1959 and 1978 the amount of acetone derived from IPA dehydrogenation declined from 80 percent to 34 percent, and the amount of IPA used for this purpose declined from 47 percent in 1978 to 12 percent in 1990.

Fig. 22.27. Propylene oxide–*tert*-butyl hydroperoxide process. (Encyclopedia of Chemical Technology, *Kirk and Othmer, 3rd ed., Vol. 19, p. 259, 1980. Copyright by John Wiley & Sons, Inc. and used by permission of the copyright owner.*)

Fig. 22.28. Isopropyl alcohol manufacture. (*Reproduced from* Pet. Ref., ***38****, no. 11, 264, 1959; copyright 1959 by Gulf Publishing Co. and used by permission of the copyright owner.*)

Isopropyl alcohol also can be made by hydrogenation of acetone, but the large capacity for IPA by indirect hydration is a disincentive for this application. In the early 1990s there were idled IPA-to-acetone plants, with little likelihood of their being reactivated. A hydrogen peroxide process based on oxidation of IPA with molecular oxygen also produces acetone.

$$H_2C{=}CHCH_3 + H_2SO_4 \rightarrow (CH_3)_2CHOSO_3H$$

$$(CH_3)_2CHOSO_3H + H_2O \rightarrow (CH_3)_2CHOH + H_2SO_4$$

Direct hydration of propylene in a vapor-phase, catalytic process also is commercially practiced. This is similar to hydration of ethylene to make ethanol. Relative to the sulfuric acid–mediated process, it offers the advantage of decreased corrosion. However, it suffers from a requirement for a pure propylene feed, whereas the former process can be used with a dilute, refinery stream.

Isopropyl alcohol is an excellent solvent with a blend of polar, nonpolar, and hydrogen-bonding character that makes it useful in a broad spectrum of applications. Its moderate volatility makes it convenient for uses involving evaporation or recovery by distillation. Thus, it is no surprise that much of IPA's consumption is for solvent uses. In the United States in 1990 about 1.4 billion lb of IPA were consumed. Major uses were: coating solvent—20 percent; processing solvent—15 percent; household and personal products—12 percent; and pharmaceutical—12 percent. The total demand in 1990 was significantly less than the 1.9 billion lb recorded for 1978. This downward trend is not unique to IPA and is primarily a result of regulatory pressure in the United States to decrease emissions of volatile organic compounds (VOCs) in coating and other applications.

Cumene

Cumene manufacture consumed about 10 percent (2 billion lb) of the propylene used for chemicals in the United States in 1988. It is prepared in near-stoichiometric yield from propylene and benzene with acidic catalysts (scheme below). Many catalysts have been used commercially, but most cumene is made using a "solid phosphoric acid" catalyst. Cumene is used almost exclusively as feed to the cumene oxidation process, which has phenol and acetone as its coproducts.

$$C_6H_6 + H_2C{=}CHCH_3 \rightarrow C_6H_5CH(CH_3)_2$$

Acetone. Acetone in commerce is derived mostly from cumene oxidation. This is a two-step process involving oxidation of cumene to the hydroperoxide followed by decomposition to acetone and phenol:

$$C_6H_5CH(CH_3)_2 + O_2 \rightarrow C_6H_5C(CH_3)_2OOH$$

$$C_6H_5C(CH_3)_2OOH \rightarrow C_6H_5OH + H_3CCOCH_3$$

The 1990 U.S. use of acetone was estimated to be 2.6 billion lb. This is approximately the same as the capacity that is limited by phenol demand. Its major uses are: methacrylic acid and esters—40 percent; solvent—20 percent; bisphenol-A—13 percent; and methyl isobutyl ketone—10 percent.

Oxo Chemicals

About 8 percent of the propylene converted into chemicals is used to make oxo alcohols such as 1-butanol and 2-ethylhexanol, which are called oxo alcohols because they are derived from olefins by the oxo process, which converts them to aldehydes. (The oxo process was described earlier in this chapter.)

Butyl Alcohols and Aldehydes. Hydroformylation of propylene gives a mixture of *n*-butyraldehyde and isobutyraldehyde. This mixture is formed approximately in the ratio of 2:1 from the high pressure, cobalt-catalyzed oxo process. There has always been a much greater demand for the linear *n*-butyraldehyde than the iso product, so it has been necessary to find uses for the latter.

The low pressure, rhodium-catalyzed oxo process has made the product mix conform to the relative demand for the two aldehydes. This process gives a 10:1 ratio of *n*-butyraldehyde to isobutyraldehyde.

Each aldehyde can be hydrogenated to the corresponding alcohol for use as a solvent or an intermediate for plasticizers and resins. *n*-Butyraldehyde also is converted to 2-ethylhexanol by sequential condensation and hydrogenation. 2-Ethylhexanol is used to make the phthalate ester, which finds wide use as a plasticizer of poly(vinyl) chloride.

Propylene Oligomers, Dodecene, and Nonene

The manufacturing processes for these materials are very similar to the one for cumene. The flow sheet in Fig. 22.29 specifically illustrates the manufacture of dodecene (propylene tetramer). When nonene is the desired product, additional fractionation is required, the extent of which is determined by product specifications.

In the reactor portion of this process, the olefin stock is mixed with benzene (for cumene) or recycle lights (for tetramer). The resulting charge is pumped to the reaction chamber. The catalyst, solid phosphoric acid, is maintained in separate beds in the reactor. Suitable propane quench is provided between beds for temperature control purposes because the reaction is exothermic.

Dodecene is an intermediate for surfactants, mainly through two routes. One, the larger user, produces dodecylylbenzene sulfonate for anionic detergents. The other goes through the oxo process to tridecyl alcohol, which then is converted into a nonionic detergent by the addition of alkylene oxides.

Nonene has two major outlets, the larger being the oxo production of decyl alcohol, which is used in the manufacture of esters, and so forth, for plasticizers. The other significant use for nonene is in the manufacture of nonylphenol, an intermediate for the important series of ethoxylated nonylphenol nonionic surfactants.

Acrylic Acid and Esters

Acetylene once was the raw material for commercial production of acrylic acid and esters, but in 1970 production of acrylic acid by oxidation of propylene was first practiced commercially. In a few years, the new process had essentially replaced the old. In 1988 about 3 percent (600 million lb) of the propylene used as a raw material for chemicals was converted to acrylic acid and acrylates.

The oxidation of propylene is carried out in two stages (Fig. 22.30). Acrolein exiting the first-stage converter can be isolated, or it can be further oxidized to acrylic acid in the second converter. The process is operated with two reaction stages to allow optimum

Fig. 22.29. Propylene tetramer process. (Pet Ref., *36, no. 11, 278, 1957; copyright 1957 by Gulf Publishing Co.*).

Fig. 22.30. Acrylic acid from oxidation of propylene. (Encyclopedia of Chemical Technology, *Kirk and Othmer, 3rd ed., Vol. 1, p. 339, 1980. Copyright John Wiley & Sons, Inc. and used by permission of the copyright owner.*)

catalyst and process conditions for each step.

$$H_2C{=}CHCH_3 + O_2 \rightarrow H_2C{=}CHCHO + H_2O$$

$$H_2C{=}CHCHO + \tfrac{1}{2}O_2 \rightarrow H_2C{=}CHCO_2H$$

$$H_2C{=}CHCO_2H + ROH \rightarrow H_2C{=}CHCO_2R + H_2O$$

Acrolein is very reactive and has some use as a chemical intermediate, as well as direct use as an aquatic pest control agent. Most acrolein is converted without isolation to acrylic acid.

Acrylic acid is almost exclusively used directly, or after conversion to an ester, as a monomer. Acrylate esters are produced by normal esterification processes. However, in dealing with acrylic acid, acrolein, or acrylates, unusual care must be taken to minimize losses due to polymerization and other side reactions such as additions of water, acids, or alcohols across the reactive double bond. Polyacrylic acids find use in superabsorbers, dispersants, and water treatment. The polyesters are used in surface coatings, textile fibers, adhesives, and various other applications.

Glycerin

Glycerine is prepared both by synthetic routes based on propylene and by a semisynthetic route as a by-product of fat hydrolysis in soap manufacture. (See Fig. 22.31.) Before 1949 all glycerin was obtained from hydrolysis of fatty triglycerides. Since that time, the entirely synthetic approaches based on propylene chemistry have been practiced but have never displaced fat-derived glycerine from commerce. In 1990, U.S. capacity for the two routes was 1:2 synthetic:semisynthetic.

In 1990, glycerin consumption in the U.S. was about 310 million lb versus a production capacity 50 percent greater than this amount. Growth is about 2 percent/year and is expected to remain at that level. More than 50 percent of U.S. usage of glycerin is in foods, pharmaceuticals, cosmetics, and similar applications. This reflects its extremely low toxicity, sweet taste, and moisturizing and lubricating properties. Chemical uses for glycerin include use as a "starter" alcohol for polyols made by alkoxylation with ethylene oxide and propylene oxide, and as raw material for alkyd polymers, plasticizers, and explosives.

Glycerin by the Epichlorohydrin Process. In the epichlorohydin process, synthetic glycerin is produced in three successive operations, the end products of which are allyl chloride, epichlorohydrin, and finished glycerin, respectively. A portion of the allyl chloride is used to manufacture allyl alcohol, and a portion of the epichlorohydrin is used in the manufacture of epoxy resins.

Fig. 22.31. Routes for the manufacture of glycerine. (Encyclopedia of Chemical Technology, *Kirk and Othmer, 3rd ed., Vol. 11, p. 923, 1980. Copyright John Wiley & Sons, Inc. and used by permission of the copyright owner.*)

The key reaction in this process is the hot chlorination of propylene, which fairly selectively gives substitution rather than the addition reactions. In this chlorination step, fresh propylene first is mixed with recycle propylene. This mixture is dried over a desiccant, heated to 650 to 700°F, and then mixed with chlorine (C_3H_6-to-Cl_2 ratio is 4:1) and fed to a simple steel tube adiabatic reactor. The effluent gases (950°F) are cooled quickly to 120°F and fractionated. The yield of allyl chloride is 80 to 85 percent.

Hypochlorous acid then is reacted with the allyl chloride at 85 to 100°F to form a mixture of dichlorohydrins. The reactor effluent is separated, the aqueous phase is returned to makeup hypochlorous acid, and the nonaqueous phase containing the dichlorohydrins is reacted with a lime slurry to form epichlorohydrin which is steam-distilled out and given a finishing distillation.

Glycerin is formed by the hydrolysis of epichlorohydrin with 10 percent caustic. Crude glycerin is separated from this reaction mass by multiple-effect evaporation to remove salt and most of the water. A final vacuum distillation yields a 99+ percent product.

A large amount of the epichlorohydrin intermediate is purified and used in the manufacture of unmodified epoxy resins.

CHEMICALS DERIVED FROM BUTANES AND BUTYLENES

Saturated four-carbon hydrocarbons (butanes) occur in natural petroleum products such as

crude oil and the heavy vapors in wet natural gas. The saturated C_4's also are produced from other hydrocarbons during the various petroleum refining processes. The butylenes—unsaturated C_4's—do not occur in nature, but are derived from butanes or other hydro-carbons either deliberately or as by-products. The complex interrelationships of C_4 hydrocarbons, including their production and use, are described in Fig. 22.32.

The chemical uses of the C_4 hydrocarbons still account for only a small fraction of the available material. The major chemical use by far is in the manufacture of methyl *t*-butyl ether (MTBE) for gasoline, which consumed almost 5 billion lb in 1988. Other chemical uses consumed more than 2 billion lb of butanes and almost 3 billion lb of butylene and butadiene fractions.

To put the volume of C_4's used in chemical manufacture in perspective with the amount used for fuel, one finds that approximately 6 percent of the butanes and about 11 percent of the butylenes were used as chemical raw materials. The trends that affect availability of C_4 hydrocarbons for chemical and energy end uses are determined by the natural gas processors, petroleum refiners, and, to a growing extent, ethylene manufacture. Changes in technology and in the availability of optimum feedstocks have far-reaching effects on the entire product mix. For example, when the availability of LPG and ethane for ethylene manufacture has decreased, *n*-butane and the higher crude cuts have been used, and the proportion of by-product butadiene has increased.

The spectrum of products which can be derived from the four-carbon hydrocarbons is shown in Fig. 22.33. Several of these also can be produced from other raw materials and are described elsewhere in this chapter.

n-Butane Derivatives

n-Butane can be obtained from natural gas and from refinery hydrocracker streams. Most of the *n*-butane goes into fuel additive uses. The major chemical use is as a feedstock for ethylene production by cracking. The other important chemical uses for butane are in oxidation to acetic acid and in the production of maleic anhydride. In the past, butane also was the main feedstock for the production of butadiene by dehydrogenation, but it has been replaced by coproduct butadiene obtained from ethylene production.

Fig. 22.32. Origins, interrelationships, and end uses of C_4 hydrocarbons. (*Reproduced from* Chemical Economics Handbook, *p. 300.200A. March 1980; Stanford Research Institute, Menlo Park, CA, 1980.*)

Fig. 22.33. Chemicals from butanes, butylenes, LPG, and higher aliphatic hydrocarbons.

Ethylene. The largest potential chemical market for *n*-butane is in steam cracking to ethylene and coproducts. *n*-Butane is a supplemental feedstock for olefin plants and has accounted for 1 to 4 percent of total ethylene production for most years since 1970. It can be used at up to 10 to 15 percent of the total feed in ethane/propane crackers with no major modifications. *n*-Butane also can be used as a supplemental feed at as high as 20 to 30 percent in heavy naphtha crackers. The consumption of C_4's has fluctuated considerably from year to year since 1970, depending on the relative price of butane and other feedstocks. The yield of ethylene is only 36 to 40 percent, with the other products including methane, propylene, ethane, butadiene, acetylene, and butylenes. About 1 to 2 billion lb of butane are consumed annually to produce ethylene.

Acetic Acid. Acetic acid is the most important carboxylic acid produced industrially. The annual production in the United States in 1988 was almost 3.5 billion lb. As with many compounds produced on a large scale, acetic acid has several different commercial processes. The carbonylation of methanol is rapidly becoming the dominant route. (This process was described earlier in this chapter in the methanol section.) The oxidation of acetaldehyde once also was very important, but it has been largely replaced by the methanol-based route. The third commercial method is the oxidation of butane.

The liquid-phase oxidation (LPO) of light saturated hydrocarbons yields acetic acid and a spectrum of coproduct acids, ketones, and esters. Although propane and pentanes have been used, *n*-butane is the most common feedstock because it ideally yields two moles of acetic acid. The catalytic liquid phase oxidation process consumes more than 500 million lb of *n*-butane to produce about 500 million lb of acetic acid, 70 million lb of methyl ethyl ketone, and smaller amounts of vinyl acetate and formic acid. The process employs a liquid-phase, high-pressure (850 psi), 160 to 180°C oxidation, using acetic acid as a diluent and a cobalt or manganese acetate catalyst.

Figure 22.34 shows the flowsheet for this process. From the reactor, the product

Fig. 22.34. Oxidation of butane. (Pet. Ref. *38, no. 11, 234, 1959; copyright 1959 by Gulf Publishing Co.*)

mixture is passed through coolers, and then through separators where the dissolved gases are released. The major components of the oxidized crude are acetic acid, methyl ethyl ketone, and various alcohols and acids. The initial reaction involves formation of butane-2 hydroperoxide, which is not isolated. Further oxidation and decomposition of the resultant radicals produce acetic acid.

Concurrently, the hydroperoxide may be converted to methyl ethyl ketone (MEK). If the initial radical attack is at the primary rather than the secondary carbon, the process makes propionic and formic acids. Reaction conditions can be changed to produce more MEK at the expense of some acetic acid. The maximum acetic acid/MEK ratio is 6.5–7 on a weight basis. If ethyl acetate is also formed, the ratio can go down to acetic acid/(ethyl acetate + MEK) of 3.6–4, with MEK being about 55 percent of the by-product.

A portion of the acetic acid, which is the major product, can be converted in a separate unit to acetic anhydride. Acetic anhydride may be produced from acetic acid, acetone, or acetaldehyde. With both acetic acid and acetone the initial product is ketene. The ketene is highly reactive and reacts readily with acetic acid to form acetic anhydride. This all takes place at 700 to 800°C in the presence of a triethyl phosphate catalyst. With acetic acid, the reactions are:

$$CH_3COOH \rightarrow CH_2{=}C{=}O + H_2O$$

$$CH_3COOH + CH_2{=}C{=}O \rightarrow (CH_3CO)_2O$$

Acetic anhydride is used to make acetic acid esters. It is especially effective in difficult acetylations, such as in the manufacture of aspirin and cellulose acetate.

Maleic Anhydride. Maleic anhydride is one of the fastest-growing chemical end-uses for butane. About half of the maleic anhydride produced goes into the manufacture of unsaturated polyester resins, used primarily in fiber-reinforced plastics for construction, marine, and transportation industries. It also is used to make lube oil additives, alkyd resins, fumaric and malic acids, copolymers, and agricultural chemicals.

Essentially all maleic anhydride is manufactured by the catalytic vapor-phase oxidation of hydrocarbons. Prior to 1975, benzene was the feedstock of choice. By the early 1980s, however, many producers had switched to *n*-butane for economic and environmental reasons. The oxidation reaction to produce maleic acid from *n*-butane is as follows:

n-Butane is oxidized by air at about 500°C over an iron–vanadium–phosphorus catalyst. The yield is in the range of 40 to 50 percent on butane.

Isobutanes

t-*Butyl Alcohol/Propylene Oxide.* The main use for isobutane is in the peroxidation of propylene with *t*-butyl hydroperoxide. The feedstocks are propylene and isobutane, and the process is similar to the PO/styrene plant (see the benzene derivatives section). In the two-stage conversion route, oxidation of isobutane with air yields a mixture of *t*-butyl hydroperoxide and *t*-butanol in a liquid-phase reaction at 135 to 144°C. After separation of products, a molybdenum-catalyzed reaction of the hydroperoxide with propylene at 110°C yields propylene oxide (PO) and *t*-butyl alcohol (TBA).

$$\underset{\text{isobutane}}{2(CH_3)_3CH} + 3/2\,O_2 \longrightarrow \underset{t\text{-butyl hydroperoxide}}{(CH_3)_3C{-}O{-}OH} + \underset{t\text{-butyl alcohol}}{(CH_3)_3COH}$$

$$\underset{\text{propylene}}{CH_2{=}CHCH_3} + \underset{t\text{-butyl hydroperoxide}}{(CH_3)_3C{-}O{-}OH} \longrightarrow \underset{\text{propylene oxide}}{H_2C{-}O{-}CHCH_3} + \underset{t\text{-butyl alcohol}}{(CH_3)_3COH}$$

The ratio of TBA to PO, and thus the isobutane requirement, can be adjusted from approximately 2:1 to 3:1. The volume of propylene oxide produced by this process is almost 900 million lb. At this time, there is only one domestic propylene oxide producer not using this process.

When this process first was introduced, the TBA was of low value, so the TBA/PO ratio was kept to a minimum. Since that time, valuable uses have been found for the TBA. *t*-Butanol is used as an octane enhancer in gasoline, competing with methanol, ethanol, and MTBE. Dehydration of TBA yields high-purity isobutylene, which is converted primarily into MTBE, butyl elastomers, polyisobutylenes, and a variety of substituted phenols and cresols, mercaptans, and amines. More than 450 million lb of isobutylene is produced from TBA.

Other Isobutane-Based Chemicals. Isobutane can be directly dehydrogenated to isobutylene by a modification of the Houdry process. This can then be converted to MTBE. The estimated use is about 1 billion lbs of isobutane. Because of their inertness and higher vapor pressures, high-purity propane and butanes have become the important substitutes for fluorocarbons as aerosol propellants. Isobutane also can be used as a solvent in polymer processing, and as a blowing agent for foamed polystyrene.

Butylenes

Butylenes are four-carbon monoolefins that are produced by various hydrocarbon processes, principally catalytic cracking at refineries and steam cracking at olefins plants. These processes yield isomeric mixtures of 1-butene, *cis*- and *trans* butene-2, and isobutylene. Derivatives of butylenes range from polygas chemicals and methyl *t*-butyl ether, where crude butylenes streams may be used, to polybutene-1 and LLDPE, which require high-purity 1-butene. In 1989, the estimated consumption of butylenes (in billions of pounds) was: alkylation, 29.0; MTBE, 4.7; polygas and fuel uses, 1.0; isobutylene derivatives, 1.2; and *n*-butene derivatives, 1.1.

The major chemical uses for *n*-butylenes are *sec*-butyl alcohol (and MEK), butadiene, butene-1, heptenes, and octenes. In 1978, butadiene accounted for almost 70 percent of the demand, but less than ten years later, its volume was only one-tenth as large. This precipitous drop was caused by the move to ethylene coproduct as the preferred source of butadiene. At the present time *sec*-butanol and 1-butene are the largest chemical end-uses of butylene, consuming about two-thirds of the total.

sec-*Butanol and Methyl Ethyl Ketone.* The largest single end-product chemical use for *n*-butenes is in the manufacture of *sec*-butanol. A refinery butanes–butylenes stream, usually rich in butene-2, is contacted with 80 percent sulfuric acid to produce the *sec*-butyl hydrogen sulfate. Dilution with water and steam stripping produce the alcohol. Because the only significant use for the alcohol is the production of methyl ethyl ketone (MEK), the processes usually are integrated as shown in the flowsheet in Fig. 22.35. MEK is obtained in high yields at catalytic dehydration of the alcohol at 400 to 500°C.

Solvent applications account for almost 95 percent of all MEK consumption; the rest goes to chemical uses such as MEK peroxide and methyl ethyl ketoxime. The solvent applications include surface coatings, adhesives, lube oil dewaxing, magnetic tape manufacture, and printing inks. Production of MEK in 1988 totaled almost 480 million lb.

1-Butene. The amount of butylenes going to 1-butene is almost as large as that going to *sec*-butanol. The demand for high-purity 1-butene grew by almost 12 percent/year during the 1980s, spurred primarily by the development of linear low-density polyethylene (LLDPE), which requires alpha-olefin comonomers. Various processes for the production of LLDPE and high-density polyethylene (HDPE) incorporate 1-butene as a comonomer. This accounts for about 70 percent of 1-butene use. The alpha-olefin

Fig. 22.35. Methyl ethyl ketone or acetone. (Pet. Ref., *38, no. 11, 272, 1959; copyright 1959 by Gulf Publishing Co.*)

comonomers control the density and physical properties of the polymer. Comonomer concentrations generally average 2 percent by weight in HDPE and 8 percent in LLDPE. About half of the 3 billion lb per year of LLDPE produced in the United States the late 1980s was based on 1-butene, with the other half split between hexene-1 and octene-1. It is expected that 1-butene will face increased competition from hexene-1 as the preferred comonomer in the future.

The remaining 30 percent of 1-butene is divided among several uses. About 10 to 15 percent of the 1-butene is polymerized in the presence of a Ziegler-type catalyst to produce polybutene-1 resin. The markets for this resin are pipe, specialty films, and polymer alloys. Approximately the same volume of 1-butene is reacted with synthesis gas in an oxo reaction to produce valeraldehydes. These C_5 aldehydes are then hydrogenated to amyl alcohols or oxidized to valeric acid. Amyl alcohols are consumed in the production of lube oil additives and amyl acetate and in solvent uses. Valeric acid goes into lubricant base stocks and specialty chemicals.

Smaller uses of 1-butene are in 1,2-butylene oxide, butyl mercaptan, and butyl phenols. Butylene oxide, produced by the chlorohydrin process, is used as a corrosion inhibitor in chlorinated solvents. Butyl mercaptan is a precursor for organophosphate herbicides, pharmaceutical intermediates, and is used as a gas odorant.

Heptenes and Octenes. Heptene and octene are oligomers produced by the polymerization of refinery streams containing C_3 and C_4 hydrocarbons. Originally these polygas units were developed to provide a source of high-octane blending components from refinery gases, but many have since been adapted to produce heptenes and octenes for chemical uses. Heptenes are used primarily to produce isooctyl alcohol which is a precursor for lube oil additives, diisooctyl phthalate (DIOP), other plasticizers, and herbicide esters. Isooctyl alcohol was once the predominant plasticizer alcohol, but development of plasticizers from competing alcohols such as 2-ethylhexanol and linear alcohols has eroded its market. Octenes are precursors of isononyl alcohol, which in turn is the raw material in the manufacture of diisononyl phthalate (DINP) plastizers. These plasticizers compete directly with dioctyl phthalate in many applications. The total U.S. consumption of butylenes for production of heptenes and octenes was approximately 240 million lb in 1989.

Butene-2. A relatively small amount of 1-butene is isomerized to butene-2. This is done because butene-2 is a better raw material for the intentional production of butadiene. Some butene-2 is used in solvent applications, and it is also the intermediate in the disproportionation process for producing propylene from ethylene.

Isobutylene

Methyl t-Butyl Ether. By far the largest use of isobutylene is in the manufacture of methyl *t*-butyl ether (MTBE). Since its introduction in 1979, the demand for MTBE as an octane improver in gasoline has grown phenomenally. By 1990, production had reached 4.9 billion lb. The use of MTBE in gasoline is expected to continue growing because of environmental pressures and the demand for more high-octane motor fuel components. The Clean Air Act of 1990 necessitates gasoline reformulation with oxygenated compounds such as MTBE and ethanol, and this could lead to a doubling of MTBE capacity in the 1990s.

MTBE is produced by reacting methanol and isobutylene under mild conditions in the presence of an acid catalyst. The isobutylene feed is either mixed butylenes, a butylenes stream from catalytic cracking, or a butylenes coproduct from ethylene production. The reaction conditions are mild enough to permit the *n*-butenes to pass through without ether formation. Figure 22.36 shows a typical two-stage process for making MTBE. The two-stage process is used in order to obtain a C_4 raffinate that is pure enough to permit recovery of the 1-butene. A potentially limiting factor in the growth of MTBE is the availability of isobutylene in the refinery streams; this probably will mean the use of more "on-purpose" isobutylene, that is, that derived from tertiary butanol or isobutane/*n*-butane.

Another approximately 1.1 billion lb of isobutylene goes into other chemical uses. These applications include polybutenes and derivatives of high-purity isobutylene such as butyl rubber, polyisobutylenes, and substituted phenols. Isobutylene is more reactive than the *n*-butenes, but many of its reactions are readily reversible under relatively mild conditions.

Polybutenes. More than 700 million lb of butylenes are consumed in the production of polybutenes. The process involves the Friedel-Crafts polymerization of desulfurized C_4 refinery streams. Although the feed is a mixed butylenes stream, the polybutene product is essentially a polymer of isobutylene, with a few *n*-butene units occurring in the polymer chain. The optimum isobutylene concentration in the feed stream is reported to be in the 20 to 25 percent range. The polymerization reaction yields products with a molecular weight range of 300 to 3000.

The lowest-molecular-weight polymers (mol wt 300–350) are used for dielectric fluids

Fig. 22.36. Two-stage MTBE process. (*Bitar, L. S., Hazbun, E. A., and Piel, W. J.,* Hydrocarbon Processing, *63, no. 10, 64, Oct. 1984. Copyright Gulf Publishing Company and reproduced by permission of the copyright owner.*)

and specialty lubricants. They comprise a very small part of the market. Polymers in the 700 to 750 range are precursors for caulks, sealants, and gasoline additives. Polymers with molecular weights of 900 to 1500 constitute the segment with by far the largest volume. They are used in the manufacture of lube oil additives via modification with phosphorus, amine, or succinimide groups. The resulting dispersants reduce oil consumption, restore compression, and improve oil viscosity. They also are used as gasoline additives and specialized plasticizers. The higher polymer fraction (mol wt 1500–3000) mainly is formulated into adhesives, caulks, and sealants. The use of polybutenes in gasoline detergents is expected to show continued growth.

Butyl Rubber. Almost two-thirds of the demand for high-purity isobutylene is for production of butyl rubber, which is produced by the cationic polymerization of high-purity isobutylene with isoprene (2–3% by weight) at low temperatures in the presence of a Friedel-Crafts catalyst. The isobutylene must be pure in order to obtain a high molecular weight product. The elastomer's outstanding property of impermeability to air and gas makes it particularly suitable for tire liners and tire bladders and valves. Other important butyl elastomers are the halogenated products, chlorobutyl and bromobutyl. Halogenated butyl rubbers, which are more compatible with other tire elastomers and can be cured faster than butyl rubber, have grown to be of more importance in recent years.

Polyisobutylenes. Polyisobutylenes are produced by the low temperature polymerization of high-purity isobutylene. The main commercial products have two molecular weight ranges: a 40,000 to 50,000 molecular weight polymer, and a polymer with molecular weight two to four times that of butyl rubber. The lower-weight polymer is used in binder systems, and in plasticizers and tackifiers for adhesives and sealants for electrical applications. The largest application for the higher-weight polymer is in the production of lube oil viscosity improvers. In 1989, about 50 million lb of high-purity isobutylene went to make polyisobutylenes.

Other Derivatives of Isobutylene. The production of many smaller-volume chemicals is based on high-purity isobutylene. The major ones are *para-t*-butylphenol, di-*t*-butyl-*p*-cresol (BHT), 2,6-di-*t*-butylphenol, *t*-butylamine, *t*-butyl mercaptan, and isobutyl aluminum compounds. Isobutylene usage by each of these six ranged from 8 and 18 million lb. The total volume of isobutylene that went for small-volume chemical production was about 100 million lb in 1989.

The substituted phenols and cresols constitute about half the total volume of this group. *para-t*-Butylphenol is produced by the alkylation of phenol with isobutylene. The principal applications for this derivative are in the manufacture of modified phenolic resins for the rubber industry and in surface coatings. Butylated hydroxytoluene (BHT) is obtained from isobutylene and *p*-cresol. Technical-grade BHT is an antioxidant for plastics and elastomers, and is a gum inhibitor in gasoline. Food-grade BHT is an antioxidant in edible oils, preserves, and many other foods. Although the volume has decreased steadily since the early 1970s, BHT remains the most important and least expensive antioxidant for food uses. 2,6-Di-*t*-butylphenol is used to produce a wide range of plastics additives, antioxidants, and gasoline additives.

t-Butylamine is formed by the reaction of isobutylene with HCN in the presence of strong sulfuric acid. The intermediate *t*-butyl formamide then is hydrolyzed to form the amine and formic acid. This amine is used mainly to synthesize sulfonamide rubber accelerator compounds. The major use of *t*-butyl mercaptan is in odorant formulations for natural gas. The distributors of natural gas inject about one pound of odorant per million cubic feet. The mercaptan also is used as a raw material for terbufos, an organophosphate insecticide used on corn. Isobutyl aluminum compounds are produced from high-purity isobutylene, hydrogen, and aluminum. Five distinctly different isobutyl

aluminum compounds are produced in the United States. The principal end use for these compounds is as polymerization cocatalysts in the manufacture of polybutadiene, polyisoprene, and polypropylene. Other minor uses for high-purity isobutylene are in the manufacture of neopentanoic acid, methallyl chloride, and miscellaneous butylated phenols and cresols.

Butadiene

In the mid-1970s there were several major processes for making butadiene in the United States: steam cracking of naphtha, catalytic dehydrogenation of *n*-butene, dehydrogenation of *n*-butane, and oxidative dehydrogenation of *n*-butene. By 1988, more than 90 percent of all the butadiene was made as a coproduct with ethylene from steam cracking, and the only "on-purpose" production came from the dehydrogenation of butylene.

Thermal cracking of hydrocarbon feedstocks in the presence of steam at 700 to 900°C produces ethylene and several coproducts, including butylenes and butadiene. The yield of butylenes varies widely (to as high as 30%), depending on the feedstock and the severity of the cracking. The yield of butadiene is particularly high with naphthas and heavier feedstocks. Most ethylene producers recover a raw C_4 stream that contains butanes, butylenes, and butadiene. The butadiene is recovered by extraction, and the raffinate (containing butanes and butylenes) is used for gasoline blending or the production of chemicals.

The dehydrogenation process feed can be refinery streams from the catalytic cracking processes. This mixed C_4 stream typically contains less than 20 percent *n*-butenes. For use in dehydrogenation, however, it should be concentrated to 80 to 95 percent. The isobutylene generally is removed first by a selective extraction–hydration process. The *n*-butenes in the raffinate then are separated from the butanes by an extractive distillation. The catalytic dehydrogenation of *n*-butenes to 1,3-butadiene is carried out in the presence of steam at high temperature (>600°C) and reduced pressure. A typical catalyst could be a chromium-promoted calcium nickel phosphate. The oxidative dehydrogenation process for butadiene reacts a mixture of *n*-butenes, compressed air, and steam over a fixed catalyst bed of tin, bismuth, and boron.

Of the 3.5 billion lb of butadiene consumed in 1988, almost two-thirds went into the production of elastomers (SBR, polybutadiene, nitrile, and polychloroprene). Adiponitrile, ABS resins, styrene–butadiene latex, styrene block copolymers, and other smaller polymer uses accounted for the remainder. The largest single use was for styrene–butadiene copolymers (SBR and latex). Most of it was made by an emulsion process using a free-radical initiator and a styrene–butadiene ratio of about 1:3. More detailed description of the rubber and polymer used can be found in Chapters 18 and 19.

Almost 250 million lb of butadiene are converted into raw materials for nylon manufacture. Butadiene reacts with hydrogen cyanide in a two-step process to give adiponitrile, which then is hydrogenated to produce hexamethylenediamine. It also can be reacted with carbon monoxide and methanol to give first a pentene ester and then the dimethyl ester of adipic acid. Hydrolysis of the ester then gives the free acid.

HIGHER ALIPHATIC HYDROCARBONS

Cyclopentadiene

Cyclopentadiene is a product of petroleum cracking. It dimerizes exothermically in a Diels–Alder reaction to dicyclopentadiene, which is a convenient form for storage and transport. Dicyclopentadiene plus cyclopentadiene demand in the United States amounted to 130 million lb in 1988.

Dicyclopentadiene can be converted back to cyclopentadiene by thermally reversing the Diels–Alder reaction. Cyclopentadiene also undergoes the Diels–Alder reaction with other olefins, and this chemistry has been used to make highly chlorinated, polycyclic hydrocarbon pesticides. These pesticides are so resistant to degradation in the biosphere,

however, that they are largely banned from use now. It also is used as a monomer and a chemical intermediate.

Isoprene

Isoprene is the basic repeating unit in natural rubber and in the naturally occurring materials known as terpenoids. It is a diene like butadiene and is useful as a building block for synthetic polymers. The most frequently used synthetic procedure for making isoprene is acid-catalyzed reaction of formaldehyde with isobutylene, giving a dioxolane intermediate that is thermally cracked to isoprene (Fig. 22.37). Isoprene also can be recovered from petroleum refinery streams. A total of 370 million lb of isoprene was used in the United States in 1988.

Isoprene is converted to elastomers such as poly(*cis*-1,4-isoprene), which is tough, elastic, and resistant to weathering and is used mainly for vehicle tires. Recently, block copolymers of isoprene with styrene have been finding use as thermoplastic elastomers and pressure-sensitive adhesives.

n-Paraffins

n-Paraffins are the unbranched fraction of hydrocarbons found in petroleum. They can be separated from the branched and aromatic hydrocarbons by a process using a shape-selective, controlled-pore-size adsorbent. In this process, the small pores exclude branched or aromatic materials and allow them to flow through a column more readily than the linear hydrocarbons, which become adsorbed into the narrow pore structure of the adsorbent. A process schematic is shown as Fig. 22.38.

Linear Olefins

Linear, internal monoolefins are produced by dehydrogenation of *n*-paraffins. Figure 22.39 shows such a process.

Primary and Secondary Higher Alcohols

Linear internal monoolefins can be oxidized to linear secondary alcohols. The alpha (terminal) olefins from ethylene oligomerization, described earlier in this chapter, can be

Fig. 22.37. Synthesis of isoprene from isobutene. (Ullmann's Encyclopedia of Industrial Chemistry, *W., Gerhartz, Ed., Vol. A14, 5th ed., p. 629, 1985. VCH Verlagsgegellschaft, Weinheim, Federal Republic of Germany* by permission.)

Fig. 22.38. Separation of high purity C_{10}–C_{24} normal paraffins from kerosene. (Hydrocarbon Processing, ***50***, *no. 11, 184, 1971; copyright 1971 by Gulf Publishing Co.*).

Fig. 22.39. Catalytic dehydrogenation process for linear internal monoolefins. (Hydrocarbon Processing, ***50***, *no. 11, 174, 1971; copyright 1971 by Gulf Publishing Co.*)

converted by oxo chemistry to alcohols having one more carbon atom. The higher alcohols from each of these sources are used for preparation of biodegradable, synthetic detergents. The alcohols provide the hydrophobic hydrocarbon group and are linked to a polar, hydrophilic group by ethoxylation, sulfation, phosphorylation, and so forth.

CHEMICALS DERIVED FROM BENZENE, TOLUENE, AND XYLENE

Until World War II, most of the commercial aromatic chemicals in the United States and other countries were derived from the coal industry. The high-temperature carbonization of coal to produce coke from the steel industry also generated a liquid stream that was rich in aromatics. Benzene, toluene, and xylene (BTX) and other aromatics could be recovered from the coal tar by extraction and distillation. The importance of this source has greatly diminished in recent years, however, and now almost all of the 6.5 billion gallons of benzene capacity worldwide, as well as that for toluene and xylene, is based on petroleum.

There are two major sources for petroleum-based aromatics, catalytic reformate and pyrolysis gasoline. The catalytic reformate is a refinery product that occurs in the catalytic reforming of naphthenes and paraffins in low-octane naphtha to produce a high-octane product. The pyrolysis gasoline stream that results from the steam cracking of hydrocarbons to produce ethylene and propylene is a very large (and growing) source of benzene and other aromatics. Because the demand for toluene is considerably less than that for benzene, some toluene is converted back to benzene by high-temperature hydrodealkylation or by catalytic disproportionation. The amount of benzene produced at any given time by these two routes depends on the relative economics of benzene and toluene. Specific process details of these production routes from petroleum can be found in Chapter 15.

The volume of aromatic chemicals used as chemical building blocks is relatively small compared to the amount that goes into fuel. About half of the benzene and more than 90 percent of the toluene and xylenes end up in the gasoline pool.

Chemicals From Benzene

Benzene is by far the most important aromatic petrochemical raw material. During 1989, some 1.8 billion gallons were consumed in the United States. This ranks it close to propylene as a chemical building block. Benzene has a broad end-use pattern. Its most important uses are for: ethylbenzene (styrene)—53 percent; cumene (phenol)—21 percent; cyclohexane (nylon)—12 percent; nitrobenzene (aniline)—5 percent; and detergent alkylate—3 percent. Other significant uses are for chlorobenzenes and maleic anhydride. Benzene is an excellent solvent, but it has been almost entirely replaced by less toxic materials. Some of the major end products are shown in Fig. 22.40.

Styrene

Styrene is produced mainly by catalytic dehydrogenation of high-purity ethylbenzene in the vapor phase. A typical integrated process scheme for the manufacture of styrene and ethylbenzene is shown in Fig. 22.17, in the earlier section on ethylbenzene. The alkylation reaction, with excess benzene, takes place in a homogeneous system with an aluminum chloride catalyst. The ethylbenzene dehydrogenation also is catalytic, using a commercially available catalyst. The fractionation train separates high-purity styrene, unconverted ethylbenzene, and minor reaction by-products such as toluene.

A smaller amount of styrene is produced as a coproduct from a propylene oxide process. In this route, ethylbenzene is oxidized to its hydroperoxide and reacted with propylene to yield propylene oxide. The coproduct methyl phenyl carbinol then is dehydrated to styrene.

The importance of styrene has increased greatly over recent decades as a result of its use as a monomer. The major uses for styrene are in plastics, rubber, latex paints and

Fig. 22.40. Chemicals derived from benzene.

coatings, synthetic rubbers, polyesters, and styrene-alkyd coatings. In these applications styrene is either polymerized to a homopolymer (55%) or used in copolymers such as acrylonitrile–butadiene–styrene (ABS) (10%), styrene–butadiene rubber (SBR) (5%), unsaturated polyester resins (5%), and other polymer applications.

Cumene (Phenol)

Cumene has become the second largest chemical end-use for benzene. It is produced by alkylating benzene with propylene at elevated temperature and pressure in the presence of an acid catalyst. U.S. production was more than 4.5 billion lb in 1989. Of this, about 96 percent then was converted to phenol.

Before 1970, there were five different processes used to make phenol in the United States: the sulfonation route, chlorobenzene hydrolysis, the Raschig process, cumene oxidation, and the benzoic acid route. By 1978, the first three processes had essentially disappeared, and 98 percent of the remaining plant capacity was based on cumene oxidation. The oxidation process is shown in Fig. 22.41.

In this process, cumene is oxidized to cumene hydroperoxide by air at about 100°C in an alkaline environment. The oxidation products are separated, and the bottoms are mixed with a small amount of acetone and sulfuric acid and held at 70 to 80°C while the hydroperoxide splits into phenol and acetone. Total domestic capacity with this process is about 3.8 billion lb/year of phenol. In the much smaller-volume benzoic acid process, toluene is air-oxidized to benzoic acid with a cobalt catalyst. The benzoic acid then is converted to phenol by an oxidative decarboxylation reaction with air at about 240°C.

The three major uses for phenol are in the manufacture of phenolic resins, bisphenol A, and caprolactam.

Phenolic Resins. Resins such as those made from phenol and formaldehyde now account

Fig. 22.41. Manufacture of phenol and acetone by oxidation of cumene. (Hydrocarbon Processing, *Nov. 1979; copyright 1979 by Gulf Publishing Co.*)

for more than one-third of the phenol consumed in the United States. They are widely used in plywood adhesives, foundry resins, thermoformed plastics, and surface coatings. A more detailed discussion of them can be found in Chapter 19.

Bisphenol. Bisphenol-A (4,4′-isopropylidenediphenol) is used mainly in the production of polycarbonates (55%) and epoxy resins (25%), two of the fastest-growing families of plastics. Other uses are in the manufacture of flame retardants such as tetrabromobisphenol-A, polysulfone resins, and polyacrylate resins. The consumption of bisphenol in the United States in 1989 topped 1.1 billion lb.

Bisphenol is obtained by the reaction of phenol and acetone. A flowsheet for its manufacture is shown in Fig. 22.42. Phenol and acetone in a molar ratio of about 3:1 are charged to an acid-resistant stirred reactor. A sulfur-containing catalyst is added, and then

Fig. 22.42. Manufacture of bisphenol A. (Pet. ref., ***38**, no. 11, 225, 1959; copyright 1959 by Gulf Publishing Co.*).

dry HCl gas is bubbled into the reaction mass. The temperature is maintained at 30 to 40°C for 8 to 12 hours. At the end of the reaction, the mixture is washed with water and treated first with enough lime to neutralize the free acid. Vacuum and heat are applied, and water and phenol are distilled separately from the mixture. The batch is finished by blowing the molten product with steam under vacuum at 150°C to remove the odor of the sulfur catalyst. The molten bisphenol is quenched in a large volume of water, filtered, and dried.

Other Phenolics. Other major derivatives of phenol include caprolactam, salicyclic acid (aspirin), alkylated phenols, and chlorinated phenols. Chemical reactions for these products are shown below.

The uses of caprolactam are described in the section on cyclohexane derivatives. Almost 60 percent of the 27 million lb of salicyclic acid produced in 1989 went to aspirin, with the remainder mainly going to salicylate esters and phenolic resins. Aspirin (acetyl salycylic acid) has long been recognized for its analgesic and antipyretic properties. It now must share this market with the competing pain relievers acetaminophen and ibuprofen.

Cyclohexane

Cyclohexane is the basic starting material for nylon fibers and resins via the intermediates adipic acid, caprolactam, and hexamethylenediamine. Of these three derivatives, adipic acid and caprolactam account for over 90 percent

of cyclohexane consumption. Cyclohexane also is used as a solvent and as a starting material for cyclohexanol and cyclohexanone.

Although cyclohexane can be recovered from natural gasoline, most is made by liquid or vapor-phase dehydrogenation of benzene. A nickel or platinum catalyst generally is used at elevated temperature and pressure.

derived mainly from cyclohexane, there are various other routes. For example, adipic acid can be made by butylene oxidation as well as by cyclohexane oxidation; HMDA can be made starting with butadiene, adipic acid, or acrylonitrile; and caprolactam also is produced from phenol. See Chapter 21 for a more detailed discussion of nylon fibers and resins.

Adipic Acid. Nylon 66, produced from adipic acid and hexamethylenediamine (HMDA), currently is the largest-volume domestic nylon. About 90 percent of all adipic acid goes to make nylon 66 fibers and resins. Although HMDA can be made from adipic acid, a major source is from adiponitrile. The commercial synthesis of adipic acid is a two-step reaction starting with either cyclohexane or phenol. In both cases, a cyclohexanone/cyclohexanol mixture is formed as an intermediate. This mixture then is catalytically oxidized with nitric acid to the adipic product. It also can be manufactured as a by-product of the caprolactam process.

Caprolactam. Essentially all caprolactam is used in the manufacture of nylon 6 fibers. This is a fast-growing nylon with applications in carpets, textiles, and tires. Caprolactam can be produced from cyclohexane, phenol, and toluene. Cyclohexanone is produced and then reacted with hydroxylamine to give an oxime. The oxime undergoes an acid-catalyzed rearrangement to give caprolactam.

Although these nylon intermediates are

Maleic Anhydride

Prior to 1975, benzene was the feedstock of choice for maleic anhydride manufacture. By the early 1980s, for economic reasons, many producers had switched to the *n*-butane process described in the previous section titled *N*-Butane Derivatives. By 1988, all of the maleic anhydride produced in the United States came from that process. However, about half of the maleic anhydride produced abroad still comes from benzene oxidation, with a small amount being recovered as a coproduct in phthalic anhydride manufacture.

Detergent Alkylate

Alkylbenzenes are major intermediates in the manufacture of synthetic detergents. If a straight-chain alkyl group is used, the resulting product is a "soft" or degradable alkylate (LAB). Using a branched chain gives a "hard" or nondegradable alkylate (BAB). Approximately 650 million lb (mainly dodecyl- and tridecyl-benzene) of LAB were produced in 1987. The production of BAB declined from

188 million lb in 1985 to less than 10 percent of that value in recent years.

The production of linear alkylbenzenes (straight-chain alkylbenzenes) involves the liquid-phase alkylation of benzene with a mixture of linear monoolefins. A typical alkylation process is the reaction with dodecene at about 50°C using an HF catalyst. The alkylate then is sulfonated to produce linear alkylbenzene sulfonate for biodegradable detergents. The manufacture of detergents is described more fully in Chapter 26.

Nitrobenzene (Aniline)

Nitrobenzene is produced by the nitration of benzene in a mixed nitrating (nitric/sulfuric) acid solution, using either the direct nitration process or adiabatic nitration. The first process involves the direct nitration of benzene in a batch or a continuous reactor. In the adiabatic process, an excess of benzene and a mixture of nitric and sulfuric acid are added to a series of reactors. Product nitrobenzene and dilute sulfuric acid are separated, water is flashed off under vacuum, and the sulfuric acid is recycled.

Of the almost 1200 million lb of nitrobenzene produced in 1989, about 98 percent went into the production of aniline. Consequently many nitrobenzene plants are integrated with facilities for aniline production. One of the smaller uses for nitrobenzene is the production of the pain-reliever acetaminophen.

In 1989, 1 billion lb of aniline was produced in the United States. The most common route involves the hydrogenation of nitrobenzene in either the liquid or the vapor phase, over a copper–silica, nickel, or platinum–palladium catalyst. A typical aniline plant is illustrated in Fig. 22.43.

Smaller amounts of aniline are made by two other methods. In the first, nitrobenzene is reduced by reaction with scrap iron in the presence of a hydrochloric acid catalyst. The iron is oxidized to the ferrous state, and the coproduct aniline is separated. This route accounts for less than 5 percent of the current aniline production. The other process avoids nitrobenzene entirely and involves the vapor-phase ammonolysis of phenol, using an alumina catalyst. Aniline is formed with diphenylamine as a by-product. About 20 percent of the aniline is produced by this route.

Aniline is consumed as a raw material in the manufacture of a number of chemicals. The approximate distribution to major users is: *p,p*-methylene diphenyl diisocyanate (MDI)—65 percent; rubber-processing chemicals—15 percent; herbicides—5 percent; dyes and pigments—4 percent; specialty fibers—2 percent. Other uses are in pharmaceuticals and photo chemicals. Principal growth is occurring from demand for MDI and the small, rapidly growing specialty fibers.

Diisocyanates (MDI). The first step in the production of MDI (and "polymeric MDI")

Fig. 22.43. Manufacture of aniline. (Pet. Ref., *38, no. 11, 224, 1959; copyright 1959 by Gulf Publishing Co.*)

is the condensation of aniline and formaldehyde to form diphenylmethane diamine. The reaction conditions can be varied to change the isomer distribution of the product. This is followed by phosgenation to give an aromatic isocyanate product mix that corresponds to the starting polyamine.

Typically a mixture containing MDI and its dimer, trimer, and some tetramer is produced. Pure MDI can be separated by distillation. MDI is supplied in several grades, depending on the number of reactive units (—NCO groups) per molecule. The most common grade is polymeric MDI with a functionality of 2.3 to 3.0. The grades used in rigid foam production typically contain 40 to 60 percent pure MDI, with the balance being dimer and other isomers. Pure MDI is used mainly for RIM (reaction injection molding) systems.

Rigid polyurethane foams constitute the largest single use for MDI and its polymers, with total consumption of more than 500 million lb. Typical laminate and board foams contain more than 60 percent MDI, whereas pour systems pour systems and spray systems contain somewhat less. The main applications of rigid foam are in construction and in the manufacture of refrigerators and water heaters. Smaller users are packaging, tank and pipe insulation, and transportation. A minor amount of polymeric MDI is used to make foundry sand binders. Pure MDI finds use in RIM systems, specialty coatings, thermoplastic resins, high-performance casting elastomers, and spandex fibers. A more detailed description of polyurethane polymers can be found in Chapter 19.

Other Uses for Aniline. Aniline currently is used as a raw material in most of the major groups of rubber-processing chemicals: accelerators, antioxidants and stabilizers, and anti-ozonants. The most important of these are the thiazole derivatives and substituted *para*-phenylenediamines. The demand for aniline in these rubber-processing uses is expected to grow at less than 1 percent/year. In agricultural chemicals, the major use for aniline is as a raw material in the manufacture of amide herbicides for controlling annual grasses and broadleaf weeds in various crops. More than 175 commercial dyes can be made from aniline, and many others are produced from aniline derivatives. This market, however, is not expected to show further growth. Since their introduction in 1979, the use of aniline in polaramid specialty fibers has shown rapid growth. This trend is expected to continue in the future. Among the important pharmaceutical derivatives of aniline are the sulfonamides, a group of compounds used to combat infections.

Chlorobenzenes

Of the 12 different chlorobenzenes that can result from the chlorination of benzene, only three are of commercial importance: monochlorobenzene (MCB), *o*-dichlorobenzene, and *p*-dichlorobenzene. Chlorination of benzene can be done either batchwise or continuously in the presence of a catalyst such as ferric chloride, aluminum chloride, or stannic chloride. It usually is run as a three-product process, with more than 70 percent of the product being the mono. The pure compounds

are separated from the crude by distillation and crystallization.

Production of chlorobenzenes peaked in the 1960s and has decreased by about 50 percent since them. Causes for the decline include the replacement of monochlorobenzene by cumene as the preferred raw material for phenol manufacture, the elimination of the herbicide DDT, and a significant drop in solvent use. Of the almost 400 million lb made in 1988, about 70 percent was monochlorobenzene, 13 percent was *o*-dichlorobenzene, and 17 percent was *p*-dichlorobenzene.

Monochlorobenzene. The largest use for monochlorobenzene, accounting for about 40 percent of the consumption, is in the production of chloronitrobenzenes. *p*-Nitrochlorobenzene (NCB) is converted into *p*-phenylenediamines for use as antioxidants in rubber processing. A smaller use for NCB is in the synthesis the pain reliever acetaminophen. *ortho*-Nitrochlorobenzene is a raw material for producing insecticides and several azo pigments. A large number of dyes also can be derived from either chlorobenzene or nitrochlorobenzene. About 25 percent of the monochlorobenzene is used as a solvent for pesticide formulation and in MDI processing. The manufacture of diphenyl oxide, for heat transfer fluids, and phenylphenols, for disinfectants, consumes another 22 percent. About 8 percent is used to make dichlorodiphenylsulfone, an intermediate in the manufacture of sulfone polymers.

Dichlorobenzenes. More than 80 percent of the *o*-dichlorobenzene is converted into 3,4-dichloroaniline, the raw material for several major herbicides. Most of the remainder goes to various solvent applications. The major demands for *p*-dichlorobenzene come from uses in polyphenylene sulfide resins, room deodorants, and moth-control agents. Any future growth will have to come from the phenylene sulfide resins.

Derivatives of Toluene

Although the bulk of the toluene never is isolated from the gasoline pool, approximately 5.5 billion lb of high-purity (98–100%) toluene was produced for nonfuel consumption in the United States in 1988. Approximately 50 percent of this toluene was dealkylated to benzene. The second largest end-use was as a solvent for coatings, paints, and lacquers. Also of importance was the use of toluene as an intermediate in the manufacture of other chemicals, mainly toluene diisocyanate and benzoic acid.

Toluene is converted into benzene by a catalytic hydrodealkylation process at elevated temperature and pressure. The importance of this use of toluene is influenced by the relative value and demand for benzene, as benzene from this source normally is more costly than that isolated directly from refinery reformate streams. Benzene (along with xylenes) also can be obtained by the catalytic disproportionation of toluene. The volume of toluene that finds use as a solvent is expected to show a continued decline because of regulations controlling the emissions of volatile organic compounds (VOCs).

Toluene Diisocyanate (TDI)

Toluene diisocyanate is manufactured from toluene by the route indicated in the following equations:

The synthesis of TDI begins with the nitration of toluene, using a nitric acid–sulfuric acid mixture. The nitration product typically contains at least 75 percent 2,4-dinitrotoluene with the balance mostly 2,6-dinitrotoluene. The nitration product next is catalytically reduced to tolylene diamine. Lastly, the diamine mixture is dissolved in chlorobenzenes and reacted with phosgene to produce the TDI. After phosgenation, the mixture is stripped of the solvent and separated by distillation. The final product is an 80:20 isomer mixture.

Most of the TDI is reacted with polyols to produce flexible polyurethane foams. These foams are widely used as cushioning materials in furniture, automobiles, carpets, and bedding. A small amount of TDI is used to make polyurethane coatings. Polyurethanes are discussed in detail in Chapter 19.

Benzoic Acid

Benzoic acid can be produced by the liquid phase oxidation of toluene using a catalyst such as cobalt or manganese. Domestic production of benzoic acid was about 100 million lb in 1990. Of this amount, about one-half went to make phenol or phenolic derivatives. Other uses are in the synthesis of caprolactam and terephthalic acid, and as food additive, and as a plasticizer and resin intermediate.

Benzyl Chloride

The principal method for producing benzyl chloride involves the photochlorination of toluene, followed by neutralization and distillation. About 40 percent of the benzyl chloride is used to manufacture benzyl phthalates, mainly butyl benzyl phthalate, which are widely used as plasticizers. The next most important use is as a raw material in the manufacture of benzyl alcohol, for use in photography, perfumes, and cosmetics.

Chemicals from Xylene

Xylenes are obtained mainly from petroleum reformate streams in the form of a "mixed xylenes." A typical composition of this stream is about 18 percent *p*-xylene, 40 percent *m*-xylene, 22 percent *o*-xylene, and 20 percent ethylbenzene. The major chemical uses of xylene, however, require the pure isomers. The purification process involves a number of steps. First the *o*-xylene is separated from the other aromatics by distillation, with the *meta* and *para* isomers going overhead along with the ethylbenzene. *p*-Xylene can be recovered by either adsorption or crystallization processes. The flow diagram in Fig. 22.44 depicts a two-stage crystallization process for recovery of high-purity *p*-xylene from mixed xylenes. In the adsorption process, the stream is charged to a fixed bed of molecular sieves, and the selectively adsorbed *p*-xylene is recovered by washing the bed with solvent. Because the demand for *p*-xylene is far greater than that for *m*-xylene, the raffinate usually is isomerized to form more of the *para* isomer.

Toluene also may be disproportionated to

Fig. 22.44. Two-stage crystallization process for recovery of high-purity para-xylene from mixed xylenes. (*Reproduced from* Hydrocarbon Processing, *p. 175, Nov. 1985; copyright 1985 by Gulf Publishing Co.*)

form equivalent amounts of benzene and xylenes. Although this method is not widely practiced because of cost, it does have the advantage of not producing any ethylbenzene.

The total demand for the xylenes is roughly comparable to that for toluene. In 1989, approximately 5.5 billion lb of *p*-xylene and 850 million lb of *o*-xylene were produced. The volume of *m*-xylene was considerably less. The principal uses of the three xylene isomers are the production of terephthalic acid (or dimethyl terephthalate), phthalic anhydride, and isophthalic acid, respectively.

Terephthalic Acid (Dimethyl Terephthalate)

Terephthalic acid (TPA) and dimethyl terephthlate (DMT) are precursors for polyethylene terephthalate, which in turn is used in the production of polyester fibers and film, PET bottles, and other polyester thermoplastic resins. In 1988 the total U.S. production was more than 7 billion lb. Current capacity is about evenly split between TPA and DMT. In the past, the relative ease of producing high-quality DMT gave it the largest share of the terephthalate market. The trend is now toward TPA, however, as the result of technological advances that permit better purification of TPA and the use of the acid directly in polymer formation.

One process for making TPA involves the air oxidation of a solution of *p*-xylene in acetic acid in the presence of a catalyst containing cobalt, manganese, and bromine. The liquid-phase reaction is conducted at about 200°C and 20 atmospheres pressure. The mixture is cooled to recover TPA by crystallization, dissolved in hot water, and hydrogenated to remove aldehydic by-products. Fiber-grade TPA is recovered by recrystallization.

In the process shown in Fig. 22.45, a mixture of *p*-xylene and *p*-methyl toluate is oxidized, and the oxidate is esterified. The crude ester is separated by distillation, with the fractions containing *p*-methyl toluate and terephthalic aldehyde ester being recycled in the oxidation. The crude DMT is hydrolyzed in water at 250 to 280°C to terephthalic acid.

Phthalic Anhydride

Although phthalic anhydride first was made commercially from the oxidation of naphthalene, by 1980 more than 75 percent of the production had been converted to *o*-xylene as the feedstock. This is now essentially the only major use for the *o*-xylene. The flowsheet in Fig. 22.46 shows a typical process for making phthalic anhydride. The *o*-xylene is vaporized by injection into the hot gas stream and then passes through a catalyst-filled multi-tube reactor. The crude phthalic anhydride is desublimated, and any acid present is dehydrated in the pre-decomposer vessel. The crude is finally purified in two distillations. Although the fixed-bed process currently is important, there are a number of plants in which a fluidized-bed reactor is used.

About 950 million lb of phthalic anhydride are produced annually in the United States. The major uses are in plasticizers, alkyd resins, and unsaturated polyester resins. The plasticizers are esters made by reacting two moles of an alcohol, such as 2-ethylhexanol, with one mole of phthalic anhydride. These plasticizers find major use in vinyl chloride polymers and copolymers. Alkyd resins are a type of polyester resin used in surface coatings. The most rapidly growing end use is in unsaturated polyester resins for reinforced plastics.

Isophthalic Acid

Although *m*-xylene is an abundant material, it has limited demand as a chemical raw material. The only major outlet is in the manufacture of isophthalic acid. More than 180 million pounds are expected to be produced in the United States in 1992, primarily for use in preparing alkyd resins and unsaturated polyester resins.

Naphthalene Derivatives

The use of naphthalene for production of chemicals has been declining steadily, as it is being replaced by other petroleum-based materials. Current domestic consumption is only about 200 million lb. There are two

Fig. 22.45. Fiber-grade terephthalic acid from *p*-xylene and methanol. (*Reproduced from* Hydrocarbon Processing, *p. 171, Nov. 1985; copyright 1985 by Gulf Publishing Co.*)

Fig. 22.46. Production of phthalic anhydride. (Hydrocarbon Processing, *p. 157, Nov. 1985; copyright 1985 by Gulf Publishing Co.*).

commercial processes for producing naphthalene. One is the recovery of naphthalene from coal tar, and the other involves its recovery from certain aromatized petroleum fractions. Until the end of the 1950s, the only commercial source of naphthalene in the United States was coal tar. At that time petroleum-derived naphthalene became a commercial product and was quickly established as a desirable source of raw material for phthalic anhydride because of its quality, low sulfur content, and stable supply.

The oxidation of naphthalene to phthalic anhydride has long been its principal end use. In the last 20 years, however, *o*-xylene has supplanted naphthalene as the preferred raw material for phthalic anhydride, and it now accounts for about 85 percent of the PA production. The manufacture of phthalic anhydride is described above in the section on xylene derivatives.

Presently the major end uses for naphthalene are phthalic anhydride (70%), carbaryl insecticide (15%), surfactants and dispersants (12%), and synthetic tanning agents (3%). Carbaryl (1-napthyl, *n*-methylcarbamate) is a broad-spectrum insecticide. The major surfactants and dispersants are derivatives of naphthalene sulfonates; their principal applications are as commercial wetting agents, concrete additives, and rubber dispersants, and in agricultural formulations. The synthetic tanning agent are derivatives of naphthalene sulfonic acid and formaldehyde. Other smaller uses of naphthalene are in beta-naphthol and as a moth-repelling agent.

REFERENCES

1. "Acetyls," Chem Systems, Inc., Process Evaluation/Research Planning, 88-1 (July 1990).
2. *Chemical Economics Handbook*, SRI International, Menlo Park, CA, issues 10/86, 12/86, 01/87, 11/87, 01/88, 10/88, 08/89, 09/90, 02/91.
3. *Chemical Marketing Reporter*, 1983, 1986, 1987, 1988, 1989, 1990.
4. *Encyclopedia of Chemical Technology*, 3rd ed., Kirk and Othmer (Eds.), Vol 1, pp. 179, 277, 330, 740; Vol. 4, pp. 316–325, 376; Vol. 5, p. 671; Vol. 7, p. 286; Vol. 9, pp. 338, 432; Vol. 11, pp. 921, 951; Vol. 13, p. 18; Vol. 16, p. 637; Vol. 19, pp. 198, 246; Vol. 22, p. 332; John Wiley & Sons, New York, 1980.
5. *Encyclopedia of Chemical Processing and Design*, J. McKetta and W. Cunningham (Eds.), Vol. 30, pp. 85, 86, Marcel Dekker, Inc., New York, 1978.
6. "Ethylene Glycols, Glycol Ethers and Ethanolamines," Chem Systems, Inc., PERP Report 70; PERP Review No. 79-2-1.
7. "Ethylene Oxide/Propylene Oxide," Process Economics Program Report No. 2D, SRI International (Feb. 1985).
8. "Ethylene/Propylene Production Technologies," Chem Systems, Inc., Process Evaluation/Research Planning, 88-1 (July 1990).
9. Hatch, L., and Matar, S., *From Hydrocarbons to Petrochemicals*, Gulf Publishing Company, Houston 1981.
10. *Hydrocarbon Processing, Petrochemicals Handbook*, Gulf Publishing Company, Houston 1983, 1985, 1989.
11. "Synthetic Organic Chemicals," U.S. International Trade Commission Publication 2470, (1990).
12. *Ullman's Encyclopedia of Industrial Chemistry*, W. Gerhartz et al. (Eds.), Vol. A1, pp. 31, 79, 149, 161, 177; Vol. A3, pp. 475–485; Vol. A4, pp. 432–441, 447, 463, 490–493; Vol. A5, pp. 40–43; Vol. A6, pp. 240–256, 257–309; Vol. A10, p. 35, Vol. A12, p. 477; VCH Publishers, Weinheim, Federal Republic of Germany, 1986.
13. *World Petrochemicals*, SRI International, Menlo Park, CA (Jan. 1990).
14. *Hydrocarbon Processing*, p. 19 (Apr. 1990).
15. *Chemical Week*, p. 21 (Feb. 21, 1990).
16. "Glycol Ethers on Different Track to the Marketplace," *Chemical Marketing Reporter* (Aug. 20, 1990).
17. SRI International, Process Economics Program, Report No. 4C (Mar. 1990).

*Petrochemical Handbook is a special section in (usually Nov.) *Hydrocarbon Processing* magazine. Also issued under separate cover.

23

Dye Application, Manufacture of Dye Intermediates and Dyes

R. I. Steiner* and J. D. Miskie**

INTRODUCTION

Sir Isaac Newton's discovery that white light that is passed through a prism separates into a spectrum of colors proved that light is the source of all color. Light is comprised, in part, of various wavelengths of radiant energy. The human eye, with its marvelous physiology of cones, interprets the wavelengths from 400 to 700 nanometers (a nanometer being 1×10^{-9} meter), transforming this input into the realization of color. Thus within this narrow portion of the total radiant energy of light lie all the colors that are perceived. The shorter wavelengths of light give violet and blue, the midrange gives greens and yellows, and the longer waves give orange and red. Manipulation of the wavelengths then is the reason for a given hue or shade that identifies a "*red* rose" or a "*blue* bird." The absorbance and reflectance of portions of the light wavelengths by an object create in the viewer's mind the distinct and unique color of that object.

The natural world contains a fantastic variety of color. Prehistoric humans used many naturally occurring products to color their implements, clothing, and dwellings. So natural dyes and pigments early on became a part of the human heritage. Some authorities believe that as far back as 5000 B.C. natural dyes were being used to color textiles. From some time in the prehistoric past until the mid–nineteenth century, natural dyes and pigments were collected from around the world and applied, by means of elaborate recipes and procedures, to various natural fiber substrates and leather in order to obtain colorful, durable dyeings.

Then as the major European cities began to light their streets with gaslights, a new and abundant resource became available to the chemist. That resource was coal tar, a by-product of the generation of gas from coal. Using coal tar derivatives, a young English chemist, William Perkin, in 1856 discovered the first synthetic dyestuff while trying to prepare quinine. That dye was "mauve," which dyed silk to a purple shade. This

*Vice-President, Research & Development, Crompton & Knowles Corporation.
**Manager, New Product Development, Crompton & Knowles Corporation.

discovery ushered in the modern era of organic synthetic dyes and pigments. Within a year Perkin founded a dyestuff company. The success of the new coal tar dyes led other investigators into the field, and the Swiss, French, and Germans quickly joined in the new technology and established commercial dyestuff companies.

Today's manufacturers of dyes and pigments use the precepts of those earlier discoveries, but have evolved far more sophisticated and elaborate technologies. The chemistry involved in the production of dyestuffs encompasses many different intermediates and synthesis routes to produce the complex molecules and compounds that are classified as dyes. The important chemical groupings are the azo, anthraquinone, indigoid, nitro, nitroso, oxazine, phthalocyanine, quinoline, stilbene, thiazole, thiazine, and triphenylmethane classes. Each molecule of these different classes has in common a conjugated electron system that responds to specific wavelengths of light. That portion of the molecule, referred to as the chromophore, determines in large measure what the hue or shade of the dyeing will be.

There are, of course, many varieties in each of these chemical classifications, so that the result is hundreds of individual dyes each with its own specific characteristics for hue and use on substrates. Dyers have arranged this very complex and large group of chemical products into ten categories arranged by the method of application to the fiber or substrate: acid dyes, azoic dyes, basic dyes, direct dyes, disperse dyes, reactive dyes, solvent dyes, sulfur dyes, vat dyes, and mordant dyes. It is in these categories that dyestuffs are discussed in the trade.

Colour Index,[1] a publication sponsored by the Society of Dyers and Colourists (England) and the American Association of Textile Chemists & Colorists, is the most widely recognized source for the classification of dyes. The dyes are classified by usage class and hue, and the chemical class usually is disclosed. When known, the chemical constitution of the dye is identified by a five-digit constitution number.

The two major types of colorants produced today are dyestuffs and pigments. The difference between these products is evident from the definitions: a dyestuff is normally a water-soluble or water-dispersible compound that is capable of being absorbed into the substrate, whereas a pigment is normally a water-insoluble compound that requires a binder or is entrapped within the matrix of the substrate. The majority of pigments are soluble in solvents and plastics, and both dyes and pigments impart very high tinctorial value for the amount used to color a product. The major end use for dyes is in the textile, leather, and paper industries. Pigments find major uses in paint, ink, and plastic industries.

The textile industry uses a large number of dyestuffs from each of the dye categories, the choice depending on the shade, fiber used, dyeing process, end use of the textile product, requirement for fastness, and economic considerations. To provide an understanding of the interrelationships that exist among the various dye classes and fiber types, a brief survey of the major textile fibers follows.

TEXTILE FIBERS

In this survey, commercially important textile fibers are grouped by their origin. First there are the natural fibers from plant sources, cotton and flax, and those from animal sources, wool and silk. A second group consists of those fibers that are regenerated or chemically modified natural materials—the rayon and acetate fibers. The final group consists of the synthetic fibers, which include nylon, polyester, acrylics, polyolefins, and spandex.

Natural Fibers

Cotton. Cotton fibers are comprised mainly of cellulose, a long-chain polymer in which the repeating unit is a glucose anhydride molecule connected by ether links. The polymer has primary and secondary alcohol groups uniformly interspersed throughout the length of the polymer chain. These hydroxyl

units impart high water absorption characteristics to the fiber and can act as reactive sites. The morphology of the cotton fiber is a complex series of reversing spiral fibrils. The fiber in total is a convoluted collapsed tube with a high degree of twist occurring along the length of the fiber. This stable fiber occurs in nature in lengths of $\frac{1}{2}$ to 2 inches, depending on the variety and growing conditions. The diameter ranges from 16 to 21 microns (a micron being 1×10^{-6} meter).

Flax. Flax, like cotton chemically, is a cellulosic fiber that has a higher degree of crystallinity than cotton. The morphology of flax is quite different from that of cotton. Flax fibers have a long cylindrical shape with a hollow core. The fibers range in length from $\frac{1}{2}$ to $2\frac{1}{2}$ inches, with a diameter of 12 to 16 microns. Flax staple is comprised of bundles of individual fibers. Historians believe that flax was among the first fibers, if not the first, to be used as a textile fiber. However, in recent years its commercial importance has been in decline.

Wool. Wool fibers are comprised mainly of proteins; the polypeptide polymers in wool are produced from some 20 alpha amino acids. The major chemical features of the polypeptide polymer are the amide links, which occur between the amino acids along the polymer chain, and the cystine (sulfur-to-sulfur) cross-links, which occur in a random spacing between the polymer chains. The polymer contains many amine, carboxylic acid, and amide groups, which contribute in part to the water-absorbent nature of the fiber.

The morphology of wool is complex. There is an outer covering over the fiber, the corticle. There are also overlapping scales having a ratchet configuration that causes shrinkage and felting. The coefficient of friction in wool fibers is vastly different between the tip and the root, depending on which way the scales point.

Wool fibers are not round but are oval in cross section. The corticle cells constitute the major component of the fiber, and are aligned along the axis of the fiber. There is a medulla section at the center region of the fiber. Each fiber has a bicomponent longitudinal crystalline arrangement. One side of the fiber contains alpha keratin crystalline regions, and the other contains beta keratin crystalline regions. Alpha and beta keratin have different moisture absorption characteristics, and this difference is what gives wool fibers crimp and springiness. It also is the reason why wool fibers kink in conditions of high humidity.

Wool fibers are sheared from about 30 major sheep breeds. The length of the wool fiber varies from 1 to 14 inches and is dependent on the breed, the climate, and the location on the sheep's anatomy. The fibers can be very fine to very coarse, ranging from 10 to 50 microns in diameter. The longer, coarser fibers normally are used for woolen fabrics, whereas the shorter, finer ones are used for worsted fabrics.

Silk. Silk, like wool, is a protein fiber, but of much simpler chemical and morphological makeup. It is comprised of six alpha amino acids, and is the only continuous-filament natural fiber. Historians claim that silk was discovered in China in 2640 B.C. Silk fiber is spun by the silkworm as a double strand, each part having a trilobal cross section. This configuration helps give silk its lustrous appearance. The fiber is wound from the cocoon that the silkworm spins as it prepares its chrysalis. The filaments are smooth and have no twist in their length, which can vary from 300 to 1800 yards. The diameter of silk is very fine, ranging from 2 to 5 microns. Because of the labor-intensiveness of silkworm culture and subsequent preparation of the fiber, silk remains a luxury fiber with a limited market niche.

Regenerated Fibers

Rayon. Viscose rayon, like cotton, is comprised of cellulose. In the manufacturing process wood pulp is treated with alkali and carbon disulfide to form cellulose xanthate. Subsequently, the reaction mass is forced through a spinnerette and precipitated in an acid coagulation bath as it is formed into a

continuous filament. The fiber has a round striated cross section. Rayon staple is made by "breaking" the continuous strands into staple-length fibers. Viscose rayon is conventionally produced at diameters varying from 9.3 to 43 microns.

Acetate. Triacetate and diacetate fibers are manufactured by the chemical treatment of cellulose obtained from refined wood pulp or purified cotton fluff. Most of the hydroxyl groups are acetylated (esterified) by treating the cellulose with acetic acid. This determines the chemical configuration of triacetate. Acetate or diacetate is made by the saponification of one of the acetylated groups, thus restoring a hydroxyl to each cellulosic monomeric unit. Theoretically, then, diacetate has two acetylated groups in each glucoside unit. The conversion of the hydroxyl group causes these fibers to be hydrophobic and changes their dyeing characteristics drastically from those of the normal cellulosic-containing fibers. Triacetate fibers are spun by mixing the isolated reaction product (flake) with methylene chloride and alcohol. The spinning solution (dope) is forced through a spinnerette and dry-spun into continuous fibers.

An alternate method of wet spinning is also possible. Acetate fibers are spun by mixing the isolated reaction product with acetone and water. The spinning solution is formed into fibers by evaporating the solvent and coagulating the acetate in a manner similar to that for triacetate (i.e., by the dry-spinning method).

Synthetic Fibers

Nylon. In 1939 the du Pont Company introduced the first truly synthetic textile fiber. Nylon was invented by Dr. Wallace Carothers as a result of his basic research into polymer science. Chemically, nylon is a polyamide polymer. Two major types of nylon polymer are used in textiles: type 6, 6, which is made by using hexamethylene diamine and adipic acid, and type 6, which is made by polymerizing caprolactam. Nylon fibers are made by melt-spinning the molten polymer. The result is a continuous filament fiber of indeterminate length. It is spun in many deniers, with its diameter varying from 10 to 50 microns. The cross section usually is set to be round or trilobal.

Polyester. Polyester is made by the polymerization reaction of ethylene glycol and terephthalic acid. The fibers are formed by melt spinning. Commercially introduced in 1953 by the du Pont Company, polyester fibers have high strength, good wrinkle recovery, and very low moisture absorbance. The fiber usually is spun with a round cross section.

Acrylics. The first commercial acrylic fiber, Orlon, was introduced by the du Pont Company in 1950s. Acrylics are made from the polymerization of acrylonitrile. The fibers are produced by either solvent spinning (Orlon) or wet spinning (Acrilan). In the solvent-spinning process, the polymer is dissolved in a low-boiling liquid solvent such as dimethyl formamide and extruded into a warm air chamber. In wet spinning, the polymer is dissolved in a suitable solvent, extruded into a coagulation bath, dried, crimped, and collected. Although the acrylic fibers are extruded as continuous filaments, they subsequently are cut into staple-length fibers. The niche that acrylics found in the market was as a substitute for wool or in wool blends (sweaters, blankets, etc.). The cross section of acrylic fiber varies among different manufacturers, Orlon having a dog-bone configuration and Acrilan having a lima-bean shape.

Polyolefins. Polyolefin fibers are produced from the polymerization of ethylene or propylene gas. Karl Ziegler's research work with catalysts ultimately was responsible for the development of these polymers. Guilio Natta, working with Ziegler's catalyst, was able to form crystalline polymers of high molecular weight polypropylene. The fibers made from these polymers are melt-spun. The cross sections are round, and the fibers are smooth. They have extremely low dye affinity and moisture absorbance.

TABLE 23.1 Hydrophobic/Hydrophilic Characteristics of Various Fibers

Fiber	*Moisture Content, % (at 65% RH—70°F)*	*Water Retention, % (Weight Change in Water)*	*Swelling, % (Volume Change in Water)*
Cotton	7	45	45
Wool	13–15	42	42
Viscose	13	90–100	95
Acetate	6.5	20–25	NA
Triacetate	3.2	12–18	NA
Nylon	3.0–5.0	9–12	13
Polyester	0.3–0.5	3–5	0.5
Acrylic	1.0–1.5	4.5–6	2

Spandex. The first manufactured elastic fiber was commercialized in 1958 by the du Pont Company. This specialty fiber is described as a segmented polyurethane that contains "hard" segments and "soft" segments; their ratio determines the amount of stretch built into the fiber. Spandex fibers are formed by dry spinning or solvent spinning. The continuous filaments can be coalesced multifilaments or monofilaments, depending on the manufacturer.

Because most dyeings are applied from water solutions or dispersions, the effect of water absorption by the fiber is an important criterion. Table 23.1 shows the hydrophobic/hydrophilic characteristics of various fibers. The cellulosics and natural fibers are the most hydrophilic, and polyester is the most hydrophobic. Although not listed in this table, the olefin fibers also are very hydrophobic.

DYE CLASSIFICATION

This section contains a brief discussion of each class of dyes including a definition of the class, the types of fibers best suited for each class, and the mechanism by which the dye is retained by the substrate. Regardless of the dye class or the fiber being dyed, dyeing proceeds according to the following sequence:

1. Movement of the dye to the substrate.
2. Absorption of the dye into the fiber.
3. Diffusion of the dye onto the fiber polymer.
4. Retention of the dye by one of the following mechanisms:
   - Ionic bonding
   - Covalent bonding
   - Entrapment
   - Hydrogen bonding
   - Solid solution

Acid Dyes

Acid dyes contain one or more sulfonic acid substituents or other acidic groups. The term applies to the application class rather than a chemical class, as other classes of dyes also contain acid substituents. Because of the presence of the acidic substituents in the dye molecule, acid dyes are anionic and water-soluble. An example of the class is Acid Yellow 36 (Metanil Yellow).*

Acid Yellow 36

Acid dyes are applied from acidic dye baths and are used to dye wool and nylon. The anionic, negatively charged portion of the dye

*The dye names used in this chapter are those given in the *Colour Index*.[1] Trivial names, when well-established, are given in parentheses following the *Colour Index* name.

molecule in solution aligns with a cationic positively charged site on the fiber. The ionic bonding that occurs is a result of the amino groups on the fiber forming ion pairs with, typically, sulfonic acid groups present as part of the dye molecule (fiber–NH_3^+ : $^-SO_3$–Dye)-

Traditionally, acid dyes have been categorized by dyers into four distinctive groups, based largely on their unique properties and dyeing characteristics: leveling dyes, milling dyes, supermilling dyes, and metal complex dyes. An outline description of these four types follows:

I. a. Leveling Dyes for Wool: Dyestuffs have low molecular weights; are applied from strongly acid dye baths.
   b. Leveling Dyes for Nylon: Dyestuffs have higher molecular weights; are applied from neutral or weakly acidic dye baths.
   Attributes: Level, even dyeings with overall good lightfastness but poor washfastness properties.

II. Milling Dyes: Dyestuffs have high molecular weights; are applied from weak acidic dye baths.
   Attributes: Good washfastness; dyeings lack brightness.

III. Supermilling Dyes: Dyestuffs have high molecular weights; are applied from neutral dye baths, usually with an auxiliary to help ensure level dyeings.
   Attributes: Excellent wash and lightfastness properties.

IV. Metal complex Dyes (Cobalt or Chromium)
   a. 1:1 Metal Complex: Dyestuffs have one equivalent of metal for each equivalent of dye with sulfonic groups; are applied from strongly acidic dye baths.
   b. 2:1 Metal Complex: Dyestuffs have one equivalent of metal for two equivalents of dye, but have no free sulfonic groups; are applied from a neutral or weakly acidic dye bath, usually used with an auxiliary to promote level dyeing.
   Attributes: Excellent washfastness and lightfastness; dyeings are dull.

Basic or Cationic Dyes

Basic dyes have a positive charge on the dye molecule and as such are ammonium, sulfonium, or oxonium salts. The positive charge can be pendant or delocalized where the charge resonates between two hetero-atoms connected by a conjugate carbon chain.

Basic dyes are applied from weakly acidic dyebaths and are used to dye acrylic and cationic dyeable polyester. The cationic positively charged portion of the dye molecule in solution aligns with an anionic, negative charged site on the fiber that the fiber manufacturers include in the polymer. The ionic bonding that occurs is the result of the anionic dyesite forming ion pairs with the quaternary amine group present as part of the dye molecule (Fiber–R^- : $^+NH_3$–Dye).

Basic dyes were the first dye class made synthetically; "mauve" was a basic bye. The basic dyes first were used to dye silk and wool, but they had poor fastness properties. Today basic dyes dyed on acrylics or cationic dyeable polyester have high tinctorial value; they are the brightest dyes available and have unlimited color range and good fastness properties.

Cationic dyes can be divided into three classes, as illustrated below.

1. Basic Brown 1 (Bismark Brown) is an amino-containing dye that is readily protonated under the pH 2 to 5 conditions of dyeing:

Basic Brown 1

2. Basic Violet 3 (Crystal Violet) is an example of a "classical" cationic dye in which the positive charge is delocalized by resonance and may be present at any one of the basic centers at any time. These resonance forms of almost equivalent energy are one of the reasons why Crystal Violet is among the strongest dyes known. This high color value (tinctorial strength) has important commercial interest in the hectograph copying system. In this system, Crystal Violet in a wax base is transferred to the back of a typewritten copy sheet. By using paper moistened with alcohol, more than 200 good copies may be made from the master.

Basic Violet 3

Developed in the late 1930s for acetate, methine dyes derived from Fischer's Base (2,3-dihydro-1,3,3-trimethyl-2-methylene-1H-indole) and Fischer's Aldehyde ([1,3-dihydro-1,3,3-trimethyl-2H-indol-2-xylidene]-acetaldehyde) later found use in dyeing acrylic fibers. They provide bright yellow to violet dyeing with acceptable lightfastness. Basic Red 14 and Basic Yellow 11 are examples of this type with a delocalized charge:

Basic Red 14

Basic Yellow 11

Further improvements in lightfast dyeing of acrylic fibers were provided by azocyanines, which also contain a delocalized charge. They give tinctorially strong bright shades with good lightfastness, and are exemplified by Basic Blue 54:

Basic Blue

3. Examples of the third type of cationic dye, with a localized pendant charge, are Basic Red 18 and Basic Blue 21:

Basic Red 18

Basic Blue 21

Direct Dyes

Direct dyes are anionic dyes that are substantive to cellulose when applied from an aqueous dye bath containing an electrolyte. Although anionic, they are not classificated as acid dyes because the acidic substituent is not the means of attachment to the fiber.

Direct dyes provide the simplest means of coloring cellulosic fibers, normally being applied from neutral or slightly alkaline dye baths, with the addition of sodium chloride or sodium sulfate salts. The purpose of the salt is to counteract the slight negative charge that cellulosics have in aqueous conditions because it would repel an anionic dye.

For a direct dye to be substantive to cellulose, the following criteria must apply:

1. The molecule should be capable of assuming a linear configuration.

2. The aromatic nuclei should be capable of assuming a coplanar arrangement.
3. The molecule should contain groups capable of forming hydrogen bonding.
4. The hydrogen bonding groups should be widely spaced along the molecule.
5. There should be a minimum of solubilizing groups to impart solubility, and these groups preferably should lie along one side of the molecule.

The hydrogen bonding that occurs is a result of the dye molecule aligning itself with the hydroxyl groups that are present along the length of the cellulosic chains. Direct dyes in and among themselves are categorized according to their dyeing characteristics:

Class A—good migrating and level dyeing.
Class B—poor leveling, control with salt additions.
Class C—poor leveling, control with temperature.

Direct dyes mostly belong to the dis-, tris- and polyazo classes. Other miscellaneous dyes belong to oxazine, thiazole, phthalocyanine, and stilbene classes. Direct Orange 26 and Direct Black 22 are typical direct dyes.

Direct Orange 26

Direct Black 22

In addition to their use to dye cellulosic fibers, direct dyes are important for coloring paper.

Since the early 1970s a new class of colorants called cationic direct dyes has become important for coloring paper. These dyes combine the molecular shape and geometry of direct dyes with the overall positive charge of cationic dyes. The first commercial products of this group are examplified by the blue dye based on copper phthalocyanine.[2]

Disperse Dyes

Disperse dyes are nonionic dyes with infinitely low water solubility, used to dye polyester, acetate, triacetate, and nylon fibers from aqueous dye baths where the dye particles are suspended by means of a protective colloid. The dye particles must be dispersed in a fine and uniform manner with a dispersing agent to distribute the dye evenly throughout the dye bath, to prevent filtration of the dye by the fiber being dyed, and to present a large surface area of dye particles. The dye molecules generally are small and have hydroxyl or amino groups to give finite water solubility at dyeing temperatures. Disperse dyes are retained by the fiber as a solid solution. The dye dissolves into the fiber at or near the glass transition temperature of the polymer.

Drawn from a large number of chemical groups, disperse dyes are referred to as high energy, medium energy, and low energy products.

1. High energy disperse dyes are intended for polyester fibers and are applied in pressure dyeing equipment at temperatures of 260 to

270°F or by continuous methods, that is, thermasol or thermafix dyeing methods at temperatures of 375 to 410°F. These dyes possess good sublimation fastness characteristics. Examples of high energy dyes are Disperse Blue 79 and Disperse Red 177:

Disperse Blue 79

Disperse Red 177

2. Medium energy disperse dyes cover a wide latitude of end uses and application methods, and usually are applied to polyester by atmospheric dyeing using a carrier. These dyes usually have fair sublimation characteristics. Examples of medium energy dyes are Disperse Orange 25 and Disperse Blue 27:

Disperse Orange 25

Disperse Blue 27

3. Low energy disperse dyes are intended for the dyeing of acetate, triacetate, and nylon, and represent the only practical way to dye acetate and triacetate. Also low energy dyes, generally with molecular weights between 250 and 350, are used in transfer printing. These dyes have poor sublimation characteristics. Examples of low energy dyes are Disperse Yellow 3 and Disperse Red 4:

Disperse Yellow 3

Disperse Red 4

Reactive Dyes

Reactive dyes possess a group(s) that will combine with the substrate, usually cellulose, and thus become part of the fiber. Reactive dyes generally contain labile halogen atoms that can undergo nucleophilic displacement by the hydroxyl of cellulose. It is necessary

to add alkali, such as soda ash, trisodium phosphate, or caustic, at some stage in the dyeing cycle to facilitate the condensation of the reactive dye with the fiber to form an ether linkage.

$$\text{Dye—X} + \text{Cell—OH} + \text{NaOH} \rightarrow \text{Dye—OCell} + \text{NaX} + H_2O$$

Some reactive moieties include the following:

Monochlorotriazine

Dichlorotriazine

Dichloropyrimidine

Trichloropyrimidine

Chlorodifluoropyrimidine

2,3-Dichloroquinoxaline

Another important type of reactive dye is the sulfatoethyl sulfone ("vinyl sulfone"), which involves an activated vinyl sulfone grouping that can react with a cellulose hydroxy in the presence of a base, as follows:

$$\text{Dye } SO_2CH_2CH_2OSO_3Na \rightarrow$$

$$\text{Dye } SO_2CH{=}CH_2 + \text{Cell—OH} \rightarrow \text{Dye } SO_2CH_2CH_2O \text{ Cell}$$

The newest development in reactive dye chemistry involves the combination of different reactive types such as a monochlorotriazine and a vinyl sulfone:

These bifunctional dyes, which also are termed multi-anchor dyes, are distinguished by their ability to achieve a good yield over a wide temperature range, by low salt demand, by a high degree of exhaustion and fixation, and by good chlorine resistance.

Reactive dyes are used primarily to dye cellulosics, cotton, and rayon, but also can be used to dye nylon and wool. The reaction in the latter case is with the amino group instead of hydroxyl. Fiber reactive dyes as a class are noted for their bright shades and excellent wetfastness properties.

Sulfur Dyes

Sulfur dyes are insoluble dyes that must be reduced with sodium sulfide before use. In the reduced form they are soluble and exhibit affinity for cellulose. They dye by adsorption, as do the direct dyes, but upon exposure to air they are oxidized to re-form the original insoluble dye inside the fiber. Thus, unlike the direct dyes, they become very resistant to removal by washing.

The exact constitutions of most sulfur dyes are unknown although the conditions required to reproduce given types are well established. They are fairly cheap and give dyeings of good fastness to washing, as noted above. Their brightness and fastness to bleaching often are inferior.

Vat Dyes

Vat dyes are water-insoluble organic pigments that become water-soluble when mixed with powerful reducing agents in the dyeing process. The water-soluble species subsequently is oxidized, and the pigment is reprecipitated *in situ*. The reducing operation formerly was carried out in wooden vats, giving rise to the name vat dye. Vat dyes are used to color cotton and rayon fibers.

Vat dyes are classified into two major groups:

I. Anthraquinoids: attributes are superior wash and lightfastness.
II. Indigonoids: attributes are brilliant shades, excellent wash and bleach fastness.

The application of vats can be accomplished by continuous methods or by exhaust dyeing procedures. In either case the sequence of events is as follows:

- Reduce the vat pigment to the "leuco" form.
- Apply the "leuco" form evenly to the substrate.
- Oxidize the "leuco" form to reprecipitate the pigment.
- Soap off to remove untrapped pigment and develop the shade.

Thus a vat dye is retained as an insoluble pigment, entrapped *in situ* in the interstices of the fibers.

An example of a vat dye is Vat Blue 4 (Indanthrone).

Vat Blue 4

Vat dyes are quite expensive and must be applied with care. They offer excellent fastness when properly selected and are the dyes most often used on cotton fabrics that are to be subjected to severe conditions of washing and bleaching.

It sometimes is impossible to tolerate the strongly alkaline conditions used to reduce vat dyes, for example, when dyeing fibers that are sensitive to alkali. For this reason and for added convenience, some manufacturers offer soluble vat dyes, which usually are the sodium or potassium salts of the sulfuric esters of reduced vat dyes. When applied to the fiber and subjected to an acid treatment in the presence of an oxidizing agent, they hydrolyze, reverting to the original form of the dye.

TABLE 23.2 Fiber and Dye Class Combinations

Origins	*Fiber*	*Dye Class*
Natural	Cotton	Direct/Reactive/Vat
	Wool/Silk	Acid/Reactive
Regenerated	Rayon	Direct/Reactive
	Acetate	Disperse
	Triacetate	Disperse
Synthetic	Acrylic	Basic
	Nylon	Acid/Reactive/Disperse
	Olefin	Na/Dope Dyed
	Polyester	Disperse
	Spandex	Acid/Disperse

Table 23.2 summarizes the relationships that exist between the various fibers and dye classes. The dyer will choose the dye class that is best suited for any given fiber based on the many criteria that the end use of the product demands.

THE APPLICATION OF DYES

The process of dyeing may be carried out in batches or on a continuous basis. The fiber may be dyed as stock, yarn, or fabric. However, no matter how the dyeing is done, the process always is fundamentally the same: dye must be transferred from a bath—usually aqueous—to the fiber itself. The basic operations of dyeing include: (1) preparation of the fiber; (2) preparation of the dye bath; (3) application of the dye; and (4) finishing. There are many variations of these operations, depending on the kind of dye. The dyeing process is complicated by the fact that single dyes seldom are used. The matching of a specified shade may require from two to a dozen dyes.

Fiber Preparation

Fiber preparation ordinarily involves scouring to remove foreign materials and ensure even

access to dye liquor. Some natural fibers are contaminated with fatty materials and dirt, and synthetic fibers may have been treated with spinning lubricants or sizing that must be removed. Some fibers also may require bleaching before they are ready for use.

Dye-Bath Preparation

Preparation of the dye bath may involve simply dissolving the dye in water, or it may be necessary to carry out more involved operations such as reducing the vat dyes. Wetting agents, salts, "carriers," retarders, and other dyeing assistants also may be added. "Carriers" are swelling agents that improve the dyeing rate of very hydrophobic fibers such as the polyesters. Examples are *o*-phenylphenol and biphenyl. Retarders are colorless substances that compete with dyes for dye sites or form a complex with the dye in the bath and act to slow the dyeing rate. Their use is necessary when too rapid dyeing tends to cause unevenness in the dyeings.

Finishing

The finishing steps for many dyes, such as the direct dyes, are very simple: the dyed material merely is rinsed and dried. Vat-dyed materials, on the other hand, must be rinsed to remove reducing agent, oxidized, rinsed again, and soaped before the final rinsing and drying steps are carried out. Generally, the finishing steps must fix the color (if fixation has not occurred during application) and remove any loose dye from the surface of the colored substrate. Residual dyeing assistants such as "carriers" also must be removed.

The types of textile structures that lend themselves to continuous dyeing methods are woven and tufted carpets. Continuous dyeing is designed for long runs of similar product; it is a high-output method of dye application.

The first volume-yardage continuous process was the continuous pad–steam process for vat dyes on cotton.[3] The vat dye dispersion was padded onto the cloth and dried; this was followed by passage through a reducing bath, steaming for 30 seconds, an oxidizing bath and, finally, washing. When it was discovered that disperse dyes could be thermosoled into polyesters by dry heat for 60 seconds at 400°F, this procedure was readily adapted to continuous processing. The advent of large volumes of dyed polyester–cotton-blend fabrics in the late 1960s made it possible to combine these two processes into one thermosol pad–steam system. See Fig. 23.1.

Kusters Corporation developed the first commercially successful dyeing range for the continuous dyeing of carpet. This was accomplished by a dye applicator that flooded the dye solution onto the carpet surface. Other machine companies soon followed with their own designs, for example, Fleissner, Gerber, and Artos. Today Kuster is marketing its Fluidyer, Gaston County its Compumat, and Otting its Spectrograph system, and Fleissner has a newly developed applicator. The essential techniques of these new systems are outlined in Table 23.3.

The advancing technology in continuous dyeing systems has created a need for dyes in liquid form, both dispersions and solutions. The dyes used in carpet dyeing, for the most part, are being supplied by the dye manufacturers as liquids.

Dyeing Methods/Batch

Virtually all types of textile structures are dyed by batch (or exhaust) methods of dyeing, such as stock, yarn, circular knits, warp knits, woven fabrics, garments, carpets, and so on. Batch methods include beck dyeing, jig dyeing, jet dyeing, package dyeing, skein dyeing, pad-batch, beam dyeing, and others. These methods are dictated primarily by the physical structure of the textile product and the type of fiber(s) it contains. Each of these batch methods employs a different type of machine. As an example, a circular knit fabric comprised of cotton could be dyed in a beck, whereas the same structure comprised of polyester most likely would be dyed in a high pressure jet machine, and a garment constructed from the circular knit cotton likely would be dyed in a paddle machine.

Continuous thermosol pad-steam dyeing system for continuous dyeing of polyester cotton blends. (*Courtesy E. I. duPont de Nemours & Co., Inc.*)

Fig. 23.1. Continuous thermosol pad–stream dyeing system for continuous dyeing of cotton blends. (*Courtesy E. I. du Pont de Nemours & Co., Inc.*)

TABLE 23.3 Recent Advances in Continuous Carpet Dyeing Systems

Fluidyer	It uses foam rather than liquid dye bath. Dye is forced out of a fixed slot over the width of the carpet.
Fleissner	Liquid dye solution flows over a weir onto the carpet, which passes over a small roller that opens the tuft at the dye contact point.
Spectrograph	Jet printing method; it uses liquid dye solutions. Single-heat jet printers can deliver four colors.
Compumat	It uses an applicator of four dye headers positioned above the carpet. The system is computer-controlled.

Stock dyeing often is carried out in large heated kettles made of stainless steel or other corrosion-resistant metal. These kettles can be sealed and used for dyeing at temperatures somewhat above the boiling point of water at atmospheric pressure.

Yarns are dyed in package machines. In this arrangement the yarn is wound onto perforated dye tubes and placed on spindles that are fitted into a closed kettle. The dye solution is heated and pumped through the spindle and yarn package. A cycle of inside/outside flow usually is used to provide level dyeing by equal exposure of the dye to yarns. Although the basis of package dyeing has not changed, a number of refinements have been introduced in recent years. Precision winding of the yarn has improved quality by giving a more uniform package density. Also lower-ratio dyebaths with higher flow rates have improved the energy efficiency of the newer machines.

Fabrics are dyed in machines that move them through the dye liquor either under tension (jig) or relaxed (beck). Fabrics also can be dyed in full width by winding them on a perforated beam through which hot dye liquor is pumped. This is the principle of the Burlington beam dyeing machine.

The pressure-jet dyeing machine is unique in that it has no moving parts. The cloth, in rope form, is introduced into a unidirectional liquid stream enclosed in a pipe. Liquor is pumped through a specially designed Venturi

jet imparting a driving force that moves the fabric. The two fabric ends are sewn together to form a continuous loop.

There are two major types of jet dyeing machines: the kier and the elongated type. In general, the kier uses small water volumes, whereas the elongated types use larger-volume ratios in dyeing. The kier types normally are used for more substantial fabrics, and the elongated types are suited for fine or delicate fabric styles. The Gaston County Dyeing Machine Company first introduced the jet machine in 1965. Important features in today's machines are improved corrosion-resistant alloys and the ability to operate at higher efficiencies with minimum energy consumption. The control systems have been refined; there is simultaneous loading and unloading. Larger-capacity machines also are being built; Gaston County now is producing a jet dye machine for carpet dyeing. See Fig. 23.2 (FUTURA 3000).

PRINTING

Printing is a special kind of localized dyeing that produces patterns. Four kinds of printing have long been recognized: (1) direct, (2) dyed, (3) discharge, and (4) resist. In direct printing, a thickened paste of the dye is printed on the fabric to produce a pattern. The fabric then is steamed to fix the dye and is finished by washing and drying. Dyed printing requires that the pattern be printed on the fabric with a mordant. The entire piece then is placed in a dye bath containing a mordant dye, but only the mordanted areas are dyeable. Washing then clears the dye from the unmordanted areas, leaving the pattern in color.

In discharge printing, the cloth is dyed all over and then printed with a substance that can destroy the dye by oxidation or reduction, leaving the pattern in white. When a reducing agent such as sodium hydrosulfite is used to destroy the dye, the paste may contain a reduced vat dye. Finishing the goods by oxidation and soaping then produces the pattern in color. In resist printing, certain colorless substances are printed on the fabric. The whole piece then is dyed, but the dye is repelled from the printed areas, thus producing a colored ground with the pattern in white.

Printing most often is done with copper rollers etched in the design to be printed. Printing paste is fed constantly to the roller from a trough, and a scraper clears the surface of the roller, leaving the dye paste only in the etched areas, from which it is transferred to the fabrics.

An important recent advance in the pattern-coloring of textiles is heat-transfer printing.[4] In this process, first the desired pattern is printed on paper, using inks containing sublimable disperse dyes. Next, the printed paper is placed in contact with a disperse-dyeable fabric, such as a polyester knit, and heat and pressure are applied to the back of the paper. The dyes transfer by subliming from the paper, diffusing across an air gap to the surface of the fabric, condensing, and then diffusing into the interior of the fiber.

PIGMENT DYEING AND PRINTING

Pigment dyeing and printing are processes that compete with the more conventional means of dyeing and printing described above. These processes use water-insoluble dyes or pigments that are bound to the surfaces of fabrics with resins. A paste or an emulsion, containing pigment and resin or a resin-former, is applied to the fabric. The goods then are dried and cured by heat to produce the finished dyeing or print. During the heating or cure, fabric, resin, and pigment become firmly bonded together. This method of color application is economical and produces good results. It should be noted that the pigment is confined to the surface of the fabric and can be selected without regard for fiber affinity.

NONTEXTILE USES OF DYES

Substantial quantities of dye are consumed by the paper industry. The most important

Fig. 23.2. FUTURA 3000 jet dyeing machine for carpet dyeing. (*Courtesy Gaston County Carpet Machinery Corp.*)

paper dyes are selected from the direct, acid and basic groups of textile dyes. The needs of the paper industry are such that it frequently is necessary to select and finish the dyes in a manner different from that employed for textiles. Now such dyes normally are supplied in solution for the convenience of the user.

Natural leather is a protein substance that can be dyed with acid dyes, among others. Finished leather also can be dyed or pigmented, the choice of color type depending on the nature of the finish.

Dyes for many materials such as gasoline, lacquers, inks, and varnishes are selected on the basis of solubility, as well as hue and fastness. It is necessary to modify some dyes to achieve the desired degree of solubility. For example, the solubility in alcohol of many acid dyes is greatly increased by forming the diphenylguanidine salts.

Important and growing markets for dyes are the plastic and metal industries. Dyed anodized aluminum, in particular, is finding increasing use in automotive and architectural applications.

Dyes for foods, drugs, and cosmetics must be chosen with great care to avoid toxic effects. In the United States, dyes are certified by the federal government when judged safe for such uses, and only the certified dyes may be legally used. This restriction does not apply to natural colors. The last synthetic dye to be certified for food use was introduced in the late 1960s.[5]

DYE INTERMEDIATES

The primary chemicals used for the manufacture of dyes are called intermediates, and are prepared from basic organic raw materials and by various chemical procedures. Intermediates, in turn, are converted to dyes by relatively simple chemical transformations.

The starting raw materials mainly are aromatic hydrocarbons such as benzene, toluene, naphthalene, anthracene, pyrene, phenol, pyridine, and carbazole, which, in the past, came almost exclusively from the distillation of coal tar but now are more commonly obtained from petroleum and natural gas. The term "coal-tar dyes," still widely considered synonymous with synthetic dyes, is no longer an entirely correct description.

U.S. dye manufacturers have relied to a great extent on European sources for intermediates. The dependence became even greater as major U.S. manufacturers—du Pont, Allied, and GAF—ceased production of dyes between 1974 and 1983. India and Pacific Rim countries are becoming important sources of intermediates. During 1983, the latest year for which government data are available, benzenoid intermediates imported into the United States were valued in excess of $840 million.[6]

In addition to the organic intermediates, a great variety of inorganic, "commodity" chemicals are used, including sulfuric acid, oleum, nitric acid, chlorine, bromine, caustic soda, sodium nitrite, hydrochloric acid, sodium carbonate, sodium bicarbonate, sodium chloride, sodium sulfate, sodium hydrosulfite, sodium sulfide, aluminum chloride, sodium bichromate, and manganese dioxide.

Intermediates are produced by a variety of reactions. Aromatic hydrocarbons undergo four electrophilic substitution reactions—the Friedel–Crafts reaction (alkylation and acylation), halogenation, nitration, and sulfonation—as well as oxidation and reduction. These reactions lead to substituted hydrocarbons that can be further reacted to provide intermediates for dyes. The number of dye intermediates actually or potentially available is very large, and the technology of their manufacture is an important part of industrial organic chemistry. Intermediates are used not only for dye manufacture but also for the manufacture of other important products such as pharmaceuticals.

The first substituent positioned in the aromatic ring of a starting material affects the position and ease of the second substitution and is called a directing group. There are two kinds of directing groups that lead to position isomers. One group directs the incoming group to *ortho* and *para* positions, whereas the other directs it to the *meta* position.

Common *ortho–para* directors are alkyl, alkoxy, hydroxy, amino, and halogen groups. *Meta* directors consist of nitro, sulfonic, carboxyl, cyano, amido, and carbonyl groups.

When position isomers are formed, it almost always is necessary to effect a separation because the isomers differ in properties and lead to different dyes. One of the most difficult problems in the manufacture of dye intermediates is the efficient separation of isomers. A further problem is to control the relative quantities in which isomers are formed or to find uses for all of them in the event that control is impractical.

The many reactions employed in the manufacture of dye intermediates and the delicate nature of many of the operations make it mandatory that all processes be well planned and controlled. Some of the processes—nitration, for example—are inherently dangerous if not properly run. For these reasons, close supervision by technically trained personnel is the rule in plants where intermediates are made.

In recent years, there has been an ever increasing emphasis on high standards of uniformity and quality in chemical manufacturing. At the same time, the increasing cost of labor has discouraged the use of numerous manual controls to meet these standards. Consequently, many plants have turned to partial or fully automatic control. Such practice requires advanced instrumentation to ensure that all process variables are closely controlled.

Because of the large number of compounds that are required, often in limited amounts, most dye intermediates are manufactured in batches. Some of the more fundamental processes can be run continuously, however, with a decided economic advantage. Where continuous production is not justified, the largest practical batches are made to hold the costs of operation to a minimum.

A wide variety of batch-type equipment is used for the manufacture of dye intermediates. Reaction kettles are made from wood, cast iron, stainless steel, or steel, often lined with rubber, glass (enamel), brick, or other corrosion-resistant material. Usual production sizes are 500 to 10,000 gallons, and the kettles are equipped with mechanical agitators, thermometers, condensers, and cooling or heating coils or jackets, depending on the nature of the operation. Products generally are transferred by gravity flow, pumping, or blowing. Isolation generally is achieved by plate and frame filter presses, filter nutsches, centrifugation, or continuous rotary filters. Both pressure and vacuum filters are used. Drying of dyes and intermediates is done in air or vacuum tray dryers or graining bowls. For larger volumes, spray dryers are becoming increasingly important.

Nitration

Nitration of aromatic hydrocarbons is a very important, fundamental reaction and produces useful dye intermediates unobtainable in any other way. Nitration involves the substitution of one or more ring hydrogen atoms by the nitro ($—NO_2$) group, which usually is accomplished by reaction of the aromatic compound with a mixture of nitric and sulfuric acid referred to as mixed acid.

The sulfuric acid serves as a solvent and generates the nitronium ion, which is the attacking electrophile:

$$HONO_2 + 2H_2SO_4 \longrightarrow 2HSO_4^{\ominus} + H_3O^{\oplus} + {}^{\oplus}N{\nearrow}O{\searrow}O$$

Nitronium ion

A typical mixed acid consists of 33 percent nitric acid and 67 percent sulfuric acid. Because nitric acid is a strong oxidizing agent, nitrations are carried out at low temperatures to avoid destructive side reactions. Some functional groups, such as the amino group, in the compounds to be nitrated must be protected against oxidative destruction by acetylation. The acetyl group is removed by hydrolysis when nitration is complete.

In nitrating benzene to nitrobenzene, an excess of nitric acid must be avoided, or dinitration will occur. The procedure is to run the mixed acids into the benzene, not the converse; this is a general rule when mononitration is desired. About 2500 lb of benzene is nitrated in one batch over a period of 3 to 4 hours. Heat is evolved during the reaction and is removed by means of cooling coils, or by means of cold water or brine circulated in a jacket surrounding the nitration vessel. Cast iron is used because it is not attacked by the mixed acids. The vessel is agitated to provide good contact between the two layers and to facilitate heat transfer. When the reaction is complete, the agitation is stopped, and the nitrobenzene separates as an oil over the acid. This oil is drawn off and agitated with water or dilute alkali to remove residual acid. It then may be distilled if pure nitrobenzene is required. In recent years, continuous vapor-phase nitration methods have been devised. Aqueous nitric acid is used, and the water resulting from the reaction is distilled off continuously to keep the nitric acid concentration high enough to be effective. Benzene can be nitrated continuously at about 80°C by using 61 percent aqueous nitric acid.

When benzene derivatives are nitrated, isomers of the desired product are obtained in addition to the product itself. For example, nitration of phenol by nitric acid gives *o*- and *p*-nitrophenol:

Phenol *o*-Nitrophenol *p*-Nitrophenol

The isomeric *o*- and *p*-nitrophenols can be separated by steam distillation. The *ortho* isomer is volatile in steam, but the *para* isomer is not and thus remains in the distillation vessel. Commercially, *p*- and *o*-nitrophenols are made by hydrolysis of the respective chloronitrobenzenes.

Mononitration of anthraquinone at about 50°C gives mainly 1-nitroanthraquinone. At 80 to 95°C, dinitration occurs to give a mixture of the 1,5- and 1,8-isomers. These isomers are important as starting materials for the preparation of other intermediates. In some cases, the mixture is used; in others, isomer separation is necessary.

Anthraquinone

1-Nitroanthraquinone

1,5-Dinitroanthraquinone 1,8-Dinitroanthraquinone

These nitrations are performed in cast iron or steel vessels with steel agitators. Because the starting materials are solids, they first are dissolved in sulfuric acid and then treated with the mixed acids; the products are also solids.

Reduction

The most common reduction reaction in the manufacture of dye intermediates is the conversion of a nitro compound to the corresponding amine.

Reduction of nitro compounds is accomplished by: (1) catalytic hydrogenation, (2) iron reduction (Béchamp), (3) sulfide reduction, or (4) zinc reduction in an alkaline medium. Generally, when the reaction is carried out on a large scale, the catalytic procedure is best. For small-scale batch operations, chemical reduction is preferred.

Catalytic hydrogenation requires a catalyst such as nickel, copper, platinum, molybdenum, or tungsten. These catalysts usually are supported on other materials and are especially prepared for the type of reduction to be carried out. Reduction conditions vary widely, depending on the nature of the nitro compound and the catalyst. Reduction may be carried out in solvent in the vapor phase or in the liquid phase. Aniline can be made by continuous vapor-phase reduction of nitrobenzene at 350 to 460°C at nearly atmospheric pressure. Some reductions, on the other hand, are run at 1000 to 4000 psi.

Iron reduction is employed on a large scale because of its simplicity. Iron turnings are used in an agitated aqueous system containing a small amount of acid to promote reaction. The overall reaction is illustrated by the reduction of nitrobenzene to aniline:

The nitrobenzene is placed in a reducer, which is a vertical cylindrical vessel provided with cover, steam jacket, and agitator. The iron turnings, or powder, and a small amount of hydrochloric acid are added in small portions. A brisk reaction is maintained by means of steam circulated in the jacket of the reducer or blown directly into the charge. A condenser returns to the reducer any vapors that escape. After the nitrobenzene is completely converted to aniline, a strong current of live steam is passed into the charge; a mixture of steam and aniline vapors passes to the condenser and is collected in storage tanks. The bulk of the aniline separates as a lower layer and is drawn off; the water over it still contains aniline, which must be recovered by distilling this "aniline water" again, or by extracting it with nitrobenzene. The iron sludge is washed out of the reducer through a side outlet by flushing. A reducer 6 feet in diameter and 10 feet high takes a charge of 5000 lb of nitrobenzene in one batch and requires about 10 hours for reduction. The aniline may be redistilled, a step that renders it water-white. However, iron reduction of nitro compounds is being deemphasized in favor of catalytic reduction, which is more efficient in a labor-intensive industry such as the dye industry.

Sulfide reduction employs sodium sulfide, sodium polysulfide, or sodium hydrosulfide. An important feature of this type of reducing system is its adaptability to bring about stepwise reduction of dinitro compounds. Partial reduction is illustrated with *m*-dinitrobenzene, which can be reduced to *m*-nitroaniline with sodium sulfide under controlled conditions:

m-Dinitrobenzene *m*-Nitroaniline

The sodium sulfide is dissolved in alcohol and placed in a steam-jacketed reducer; the dinitrobenzene is added either in solid form or dissolved in alcohol. The mixture is boiled for 2 hours, and then the alcohol is distilled off and collected for reuse. The *m*-nitroaniline mixed with inorganic salt remains in the reducer. The mass is agitated with water, which dissolved the salt, and then is pumped into a filter press. The press cake of *m*-nitroaniline is washed and then discharged and dried in a vacuum drier.

Zinc reduction in alkaline aqueous or alcoholic medium is especially useful to bring about bimolecular reduction. This kind of reaction is illustrated by the conversion of nitrobenzene to hydrazobenzene. Rearrangement of hydrazobenzene with acid gives benzidine, a formerly important intermediate for azo dyes that now is banned because of its carcinogenic activity.

Hydrazobenzene

Benzidine

In a similar way, *o*-nitroanisole is converted to hydrazoanisole and then to *o*-dianisidine.

o-Nitroanisole Hydrazoanisole *o*-Dianisidine

Amination

The introduction of an amino group into an aromatic nucleus by replacement of another functional group is called amination. This process is to be distinguished from reduction of a nitro group, in that one group is totally displaced by another and not simply altered in character.

An example of amination in the benzene series is the conversion of *p*-nitrochlorobenzene to *p*-nitroaniline with ammonia. This reaction may be carried out continuously with 40 percent aqueous ammonia under 200 atmospheres pressure at 235 to 240°C.

In the anthraquinone series, amination frequently is a convenient means of preparing amines. 1-Aminoanthraquinone-2-carboxylic acid can be made by reaction of 1-nitroanthraquinone-2-carboxylic acid with 15 percent aqueous ammonia at 130°C. The nitro group is displaced, not reduced. In a similar manner, 1-anthraquinonesulfonic acid can be aminated to give 1-aminoanthraquinone, or, if desired, ammonia may be replaced by methylamine to give 1-(methylamino)anthraquinone.

1-Nitroanthraquinone-2-carboxylic acid 1-Aminoanthraquinone-2-carboxylic acid

1-Anthraquinonesulfonic acid 1-Aminoanthraquinone

Sulfonation

Sulfonation of aromatic compounds is very important in the manufacture of dyes. It is used to make aromatic compounds water-soluble, as the salts of sulfonic acids generally are water-soluble.

Sulfonation involves the substitution of one or more ring hydrogen atoms by the sulfonic acid (—SO_3H) group. Sulfonation is achieved with (1) concentrated 96 to 100 percent sulfuric acid, (2) oleum (sulfuric acid plus sulfur trioxide), (3) sulfur trioxide in organic solvent or as a complex, or (4) chlorosulfonic acid.

Sulfonation of aromatic hydrocarbons is believed to proceed by attack of SO_3 as the electrophilic reagent, as follows:

$$2H_2SO_4 \rightleftharpoons H_3O^{\oplus} + HSO_4^{\ominus} + SO_3$$

In carrying out this reaction, oleum containing 8 percent free sulfur trioxide is added slowly to offset the dilution caused by the water formed in the process. The temperature is maintained at 30°C until near the end, when it is raised to 50°C. When the reaction is complete, the charge is diluted by running it into water, and the product is precipitated by adding salt. The product then is isolated by filtration as the sodium sulfonate.

In addition to sulfonation with sulfuric acid or its equivalent, amines may be sulfonated by baking the sulfates at elevated temperatures. This procedure offers the advantage of giving fewer isomers. The baking of aniline sulfate at 260 to 280°C gives a high yield of the *para* sulfonic acid, sulfanilic acid.

Indirect sulfonation may be achieved in a number of ways. Sodium bisulfite often will replace a labile functional group with the sulfonic acid group. An example is *o*-chlorobenzoic acid, which is converted to *o*-sulfobenzoic acid by aqueous sodium bisulfite:

The sulfonation of naphthalene yields a number of isomers, and the product may be controlled to some extent by the choice of agent. With any one agent, temperature and time of reaction determine the result. It rarely is possible to obtain a single isomer, but effort is directed toward forming a preponderant amount of one isomer. For example, for the monosulfonates, made by direct sulfonation with sulfuric acid, there is formed at 80°C, in 8 hours, 96 percent 1-naphthalenesulfonic acid. As the temperature is raised, correspondingly less of this isomer and more of the 2-sulfonic acid are formed. At 150°C, for example, the result is formation of 18 percent of the 1-isomer and over 80 percent of the 2-isomer.

In carrying out this sulfonation, the acid is run into the melted naphthalene to avoid disulfonation. The amount of acid added is the calculated amount for one sulfonic acid group. The water formed during the reaction retards but does not prevent it. Oleum may be added toward the end to hasten the reaction.

The sulfonation of amino and hydroxy derivatives of naphthalene usually leads to a large number of isomers. To avoid isomer formation as much as possible, naphthionic acid often is prepared by baking the sulfate of 1-naphthylamine.

By direct sulfonation of 1-naphthylamine, four of the seven possible 1-naphthylamine-sulfonic acids may be formed. The main product under proper conditions is the 1,4-isomer. Direct sulfonation of 2-naphthylamine yields primarily a mixture of 2,5- and 2,8-isomers.

Two important monosulfonic acids result ing from the sulfonation of 2-naphthol are Schaeffer's acid and Crocein acid. At 100°C, Schaeffer's acid is preponderant; at lower temperatures more Crocein acid is formed.

Schaeffer's acid

Crocein acid

By further sulfonation, two isomeric disulfonic acids are the main products. In the cold, G acid predominates, whereas at higher temperatures R acid is formed in greater amount. Both isomers are important intermediates for azo dyes.

R acid

G acid

Anthraquinone is sulfonated by suspending it in oleum containing 45 percent free sulfur trioxide and heating at 150°C for one hour. The resulting melt is run into water and neutralized with sodium hydroxide while still hot. On cooling, the sodium salt of the 2-sulfonic acid separates. Further sulfonation produces a mixture of the 2,6- and 2,7-disulfonic acids.

When anthraquinone is sulfonated in the presence of mercury sulfate, the results differ from those just described. A single sulfonic acid group enters at position 1, whereas two groups enter to form the 1,5- and 1,8-disulfonic acids. The 1,5-isomer is salted out from the more soluble 1,8-isomer after dilution of the sulfonation mass. The 1,5- and 1,8-disulfonic acids are of great importance for the manufacture of other derivatives, which can be made by replacement of the sulfonic acid groups. Examples are the chloro- and hydroxyanthraquinones.

Halogenation

Halogen atoms, usually chlorine or bromine, are introduced by the action of the elemental halogen on the aromatic hydrocarbon in the presence of a Lewis acid catalyst. The *chlorinium* or *bromonium* ion is the attacking electrophile.

Chlorination of alkylated aromatic compounds such as toluene can occur either in the aromatic ring or in the side chain. Friedel–Crafts catalysts (see above) direct the chlorine to the ring, whereas in the presence of UV light, chlorination of the side chain occurs. The latter reaction is believed to take place by a free radical mechanism.

$$Cl_2 + FeCl_3 \longrightarrow [FeCl_4]^- + \underset{\text{Chlorinium ion}}{Cl^+}$$

$$H^+ + FeCl_4^- \longrightarrow FeCl_3 + HCl$$

Chlorine is the most widely used halogen because it is comparatively economical. In addition to direct chlorination, chlorination may be achieved with thionyl chloride, phosphorus oxychloride, phosphorus pentachloride, or sulfuryl chloride. For example, Disperse Violet 28 is prepared by the chlorination of Leuco-1,4-diaminoanthraquinone with sulfuryl chloride:

Chlorobenzene is made by passing a stream of dried chlorine into benzene in the presence of ferric chloride; some *p*-dichlorobenzene is formed at the same time. To produce monochlorobenzene as the sole product, excess benzene is only partially converted, and

Disperse Violet 28

unreacted benzene is recycled. Even so, some polychlorobenzene is obtained.

Chlorination of toluene in the complete absence of iron produces side-chain chlorination products: benzyl chloride, benzal chloride, and benzotrichloride. All these compounds are valuable although the reaction is difficult to stop to produce pure intermediate products. Benzal chloride is converted to benzaldehyde with calcium carbonate in water, whereas benzotrichloride gives benzoic acid under the same conditions:

In the naphthalene series, direct chlorination seldom is used. The reaction takes place readily but leads to numerous isomers. In the anthraquinone series, both direct and indirect chlorination are employed. An example of indirect chlorination is the conversion of 1-anthraquinonesulfonic acid to the corresponding chloro compound. This reaction is carried out at approximately 100°C with sodium or potassium chlorate in hydrochloric acid.

Chlorination of aliphatic compounds is illustrated by the chlorination of acetic acid. One, two, or three of the hydrogen atoms on carbon can be replaced by direct chlorination of the warm liquid in the presence of sulfur.

$$CH_3COOH \xrightarrow{Cl_2} \underset{\text{Monochloroacetic acid}}{ClCH_2COOH} \xrightarrow{Cl_2}$$

$$\underset{\text{Dichloroacetic acid}}{Cl_2CHCOOH} \xrightarrow{Cl_2} \underset{\text{Trichloroacetic acid}}{Cl_3CCOOH}$$

Both fluorine and bromine find some use in the manufacture of dye intermediates, but high cost restricts them to applications where they offer some unique advantage over chlorine.

Alkaline Fusion

Alkaline fusion is an important procedure for the hydroxylation of aromatic compounds. In this reaction, a sulfonic acid group is replaced by a hydroxyl group. The reaction cannot be used when nitro or chloro groups are present but is applicable to amino compounds.

Alkaline fusion usually is carried out with a concentrated solution of sodium hydroxide in a cast iron pot that is equipped with a scraping agitator and is heated externally. The water is evaporated, and then the mass fuses. The reaction temperature is between 190 and 350°C, depending on the reactivity of the sulfonic acid.

Phenol can be made by fusing benzenesulfonic acid according to the following

equations:

Phenol

In a similar way, resorcinol is made by fusion of *m*-benzenedisulfonic acid:

Resorcinol

2-Naphthol is made by fusion of the corresponding sulfonic acid with sodium hydroxide. The naphtholate is treated with carbon dioxide to precipitate the naphthol, which then is purified by vacuum distillation.

2-Naphthol

1-Naphthol is not made from the sulfonic acid. Instead, 1-naphthylamine sulfate is heated with water at 200°C in a closed vessel, and on cooling, 1-naphthol crystallizes out.

For some compounds, the conditions of alkaline fusion are too severe. In such cases the reaction often may be effected in water solution under pressure. Chromotropic acid is prepared from 1-naphthol-3,6,8-trisulfonic acid in 60 percent sodium hydroxide solution under pressure:

Chromotropic acid

Sulfonated naphthylamines may be fused without destruction of the amino group. The important azo dye intermediates, H acid and J acid, are made in this way:

H acid

J acid

In the anthraquinone series, 2-anthraquinonsulfonic acid can be converted to the hydroxy compound by heating with calcium hydroxide in water. Alkaline fusion of the same sulfonic acid gives alizarin.

2-Hydroxyanthraquinone

Alizarin

Hydroxylation takes place under milder caustic conditions if the group to be displaced is very reactive, as, for example, a chlorine substituent *ortho* to an electron withdrawing group such as a nitro ($—NO_2$) group.

2-Aminophenol-4-sulfonamide

Reduction of the nitro group provides an important intermediate for premetallized dyes, 2-aminophenol-4-sulfonamide.

Oxidation

Oxidation may be effected by air in the presence of a catalyst or by a variety of chemical oxidants such as manganese dioxide and potassium permanganate.

Catalytic vapor-phase oxidation is illustrated by the conversion of naphthalene to phthalic anhydride. This reaction is carried out over a vanadium pentoxide catalyst at 450°C.

Another route to phthalic anhydride is oxidation of *o*-xylene, a product of the petroleum industry. The conditions for this reaction are similar to those for naphthalene oxidation, except that the temperature is higher here (540°C).

Aniline sulfate can be chemically oxidized with manganese dioxide in sulfuric acid. The product is *p*-benzoquinone.

p-Benzoquinone

The use of potassium bichromate in sulfuric acid effects the oxidation of 1-nitro-2-methylanthraquinone to the corresponding carboxylic acid. This reaction illustrates side-chain oxidation, an important route to carboxylic acids.

1-Nitro-2-methylanthraquinone — 1-Nitro-2-anthraquinonecarboxylic acid

Other Important Reactions

Condensation. In a restricted sense, condensation describes those reactions in which a carbon–carbon bond is formed by the elimination of a simple molecule. An example is the condensation of benzanthrone to dibenzanthrone, shown below in Fig. 23.6 and described on page 909. Dye intermediate chemical terminology uses the term "condensation" more broadly than this to describe a variety of reactions in which molecules are joined; the bonds formed are not restricted to carbon–carbon. An example is the formation of the diphenylamine derivative, commonly called Nitro Delta Acid:

Nitro delta acid

Addition. An important intermediate, cyanuric chloride, is made by the addition of cyanogen chloride to itself. The reaction is catalyzed by a small amount of free chlorine.

Cyanuric chloride

Alkylation. Alkylation refers to the introduction of an alkyl group, such as methyl, into an organic molecule. Alkylation may occur on carbon, nitrogen, oxygen, or sulfur. Of these possibilities, alkylation on nitrogen and alkylation on oxygen are most important in the manufacture of dye intermediates. A methyl group may be introduced into the amino group of aniline by heating with methyl alcohol under pressure in the presence of mineral acid:

N-Methylaniline

Phenol may be methylated with methyl sulfate in cold alkaline medium to give anisole:

Anisole

Carboxylation. The carboxylic acid group may be introduced by side-chain oxidation as described above. In addition, it may be introduced by direct action of carbon dioxide on certain compounds. When sodium phenolate is treated with carbon dioxide under pressure at about 150°C, sodium salicylate is formed. Acidification of the salicylate gives the free acid, which may be purified by vacuum distillation. 3-Hydroxy-2-naphthoic acid, an intermediate for developed azo dyes, is made from sodium 2-naphtholate.

Sodium salicylate Salicylic acid

3-Hydroxy-2-naphthoic acid

Sandmeyer Reaction. The replacement of a diazonium group by halogen, nitrile, nitro, sulfhydryl, and other groups in the presence of a cuprous salt is known as the Sandmeyer reaction. An illustration of the use of this reaction is the preparation of 2-chloro-5-nitrophenol from the corresponding aminophenol:

2-Amino-5-nitrophenol

2-Chloro-5-nitrophenol

Cyanoethylation. Reaction of a primary or a secondary aromatic amine with acrylonitrile results in *N*-cyanoethylation. The products are useful intermediates for azo and basic dyes. An example is the cyanoethylation of aniline with cupric sulfate catalyst:

N-(Cyanoethyl)aniline

Ethoxylation. Phenols and primary or secondary aromation amines can be reacted with ethylene oxide to provide 2-hydroxyethyl ethers or amines:

To favor formation of the mono-substituted aniline, a large excess of aniline is used,

whereas for the disubstituted product, 2:1 molar equivalents of aniline to ethylene oxide are used. These products and their derivatives are used for the manufacture of disperse dyes.

Bucherer Reaction. The reversible conversion of a naphthylamine to a naphthol in the presence of an aqueous sulfite or bisulfite is valuable in the synthesis of naphthalene dye intermediates. It is used in the preparation of naphthols from naphthylamine or is employed for the reverse transformation—naphthylamines from naphthols:

THE MANUFACTURE OF DYES

Earlier in the chapter it was stated that dye molecules possess a conjugated electron system. It is this system that confers on the dye molecule the ability to absorb light in the visible region of the spectrum. The main structural unit of a dye, the *chromophore*, is always unsaturated. A compound containing a chromophore is called a *chromogen*, and any substituent atom or group that increases the intensity of the color is called an *auxochrome*. An auxochrome also may serve to shift the absorption band of a chromophore to a longer wavelength, or may play a part in solubilizing the dye and attaching it to fibers.

Dyes are classified according to the type of chromophore that they contain. An example is the azo class, which is characterized by the presence of the —N=N— linkage. A typical chromogen is azobenzene:

Common auxochromes are hydroxyl, amino, and carboxyl groups.

The hue, strength, and brightness of a dye depend on the entire light-absorbing system, consisting of the chromophore and auxochromes acting together. The nature of these groups and their relative positions in the molecule must be worked out correctly to produce a dye of the desired appearance.

In general, for a given type of dye, extension of the unsaturated system and increased opportunities for resonance shift the absorption of light toward longer wavelengths. Assuming a single main absorption band, the color absorbed progresses across the visible spectrum from violet to red. As this progression occurs, the light that is not absorbed is reflected and seen by the human eye as the color complementary to that absorbed. The wavelengths absorbed, the corresponding colors, and the observed complementary colors are given below for several major hues.

Wavelength Absorbed nm	*Corresponding Color Absorbed*	*Observed Color*
430	Violet	Yellow
470	Blue	Orange
530	Green	Red
570	Yellow	Violet
600	Orange	Blue
660	Red	Green

In addition to securing the desired appearance, it is necessary in dye synthesis to provide the dye with any groups necessary to confer solubility and affinity for textile fibers. It also is important only to use color systems that have the required fastness to light and other degrading influences. Considering that all these properties must be exhibited by one compound, it can be readily understood that the development and the manufacture of dyes require a high degree of technical competence in the laboratory and in the plant.

Nitro Dyes

Dyes of this class are the *ortho-* or *para-*nitrophenols, nitronaphthols, and nitroamines. The chromophore is the nitro ($—NO_2$) group, and the auxochromes are hydroxyl and amino groups.

One of the oldest nitro dyes is picric acid (2,4,6-trinitrophenol):

Picric acid

An important dye in this class is the substituted nitrodiphenylamine, Disperse Yellow 42, which is used to dye polyester with very good lightfastness.

Disperse Yellow 42

Azo Dyes

The azo dye class represents the largest and most important group of dyes and includes many hundreds of commercial dyes of various application types. The azo chromophore ($—N{=}N—$) forms part of the conjugated system and joins two or more aromatic rings.

The principal method of forming azo dyes involves *diazotization* of primary aromatic amines, followed by *coupling* with hydroxy or amino derivatives of aromatic hydrocarbons or with certain aliphatic keto compounds. Both the aromatic amine, which is diazotized, and the compound to which it is coupled may bear a variety of substituents, such as alkyl, alkoxyl, halogen, and sulfonic acid. Because of the large number of compounds that can be combined, often in more than one sequence, the number of possible azo dyes is almost infinite.

Diazotization takes place when nitrous acid (HNO_2) reacts with a primary aromatic amino group in an acid medium. Usually sulfuric or hydrochloric acid is used, and the nitrous acid is generated from sodium nitrite.

The equation for the overall process using aniline as the amino compound is as follows:

$$C_6H_5—NH_2 + 2HCl + NaNO_2 \rightarrow$$
$$C_6H_5N^+{\equiv}NCl^- + 2H_2O + NaCl$$

Diazotization usually is carried out with excess acid to prevent partial diazotization and to inhibit secondary reactions. If the reaction is to proceed easily, the amine must be in solution, or its hydrochloride, if insoluble, must be in a fine state of subdivision. Temperatures of from 0°C to 10°C usually are employed for diazotization because diazonium salts generally are unstable. There are exceptions to this, and temperatures of 20°C or higher occasionally are preferred.

The ease of diazotization depends markedly on the basicity of the amine. Extremely weakly basic amines are diazotizable only by special methods, usually with nitrosylsulfuric acid ($NOHSO_4$).

The coupling reaction with aromatic hydroxy compounds is illustrated by benzene diazonium chloride and phenol:

Coupling to phenols, naphthols, and related hydroxyl compounds is carried out in alkaline solution. Under alkaline conditions the hydroxy compound is soluble, and coupling usually is rapid. Ordinarily the coupling must be carried out in the cold to prevent decomposition of the diazonium salt.

An example of coupling to an aromatic amine is the reaction of *p*-nitrobenzenediazonium chloride with *m*-toluidine:

Couplings to amines are carried out in acid solution in which the amine is soluble. It has been shown, however, that the free amine couples, not its hydrochloride. For this reason coupling proceeds faster near the neutral point where the equilibrium concentration of free amine is highest. It is frequently best to start the reaction at a low pH and then to raise the alkalinity slowly with sodium acetate or soda ash as the reaction proceeds.

It should be noted that in the above examples of coupling the diazonium group is shown entering the position *para* to the hydroxyl and amino groups. This specificity is characteristic of azo coupling. The diazo group does not enter at random but in certain definite positions. For the benzene derivatives, the attack is always on the position *para* to the activating group. If this position is blocked, coupling sometimes will occur in one of the *ortho* positions, but at a much slower rate.

In naphthalene derivatives, orientation of the entering diazo group is somewhat different. In *alpha*-naphthol, the attack is at position 4. If 4 is blocked, the diazo group enters at 2. In *beta*-naphthol, coupling takes place at 1, never at 3 or 4. The same rules apply to the corresponding naphthylamines.

alpha-Naphthol *beta*-Naphthol

In the naphthalene series, it is found that the presence of certain substituents, especially the sulfonic acid group, can influence the position of coupling even when not in a directly blocking position. For example, a sulfonic acid group in position 5 of *alpha*-naphthol causes the coupling to occur predominantly at position 2, rather than 4.

Certain aminonaphthol sulfonic acids couple twice. In this case, the place of entry depends on whether the coupling is performed in acid or in alkaline solution. If acid, the coupling is *ortho* to the amino group; if alkaline, it is *ortho* to the hydroxyl. In the three examples following the place of entry for acid coupling is marked X, for alkaline coupling Z.

H acid

Gamma acid

J acid

Ortho and *para* diamines in both the benzene and naphthalene series do not couple at all; only the *meta* diamines do so.

A generalized procedure for diazotization and coupling as practiced in both laboratory and plant is as follows: Sodium nitrite is added slowly with stirring to an acid solution of the amine in a wooden or brick-lined tub. The tub is not externally cooled, but ice is added directly to the amine solution to control the temperature between 0°C and 10°C. The progress of the reaction is followed by observing the disappearance of the nitrous acid generated from the sodium nitrite. This is done by spotting the reaction mass on a starch–potassium iodide test paper, which turns black when nitrous acid is present. When no more nitrous acid is consumed, the diazotization is judged to be complete. The diazonium compound is not isolated, but is run at once at a slow rate into the alkaline (or acid) solution of the intermediate with which it is to couple. After addition of the diazonium compound, the batch is stirred until coupling is complete. Then the solution of the dye is warmed, and salt (NaCl) is added and allowed to dissolve. On cooling, the dye separates and is filtered. There are a number

of variations in procedure. Some dyes are salted hot, some cold, some not at all. Filtration is carried out at a temperature best suited to isolate the dye without inclusion of impurities. As a result of salting, dyes with sulfonic acid groups always are isolated as the sodium sulfonates.

Diazotization and coupling usually are carried out as batch processes because of the limited production required and the necessity of testing the various batches of intermediates that are used. In a few cases, however, continuous diazotization and coupling are practical. When continuous processes are used, it becomes necessary to devise a means of isolating the product as rapidly as it is made. Centrifugal filters, spray dryers, and other special apparatus are suitable for this use.

Monoazo Dyes. Monoazo dyes contain one azo (—N=N—) linkage; the simplest example is aminoazobenzene. This compound is used as a dye for oils, lacquers, and stains under the name Solvent Yellow 1 (Aniline Yellow), and is also an intermediate for other dyes.

Solvent Yellow 1

To simplify this description of azo dyes, only the formulas of the dyes are given, with the place of attack of the diazonium group indicated by an asterisk on the active carbon of the coupling component. It should be understood that the diazonium compound was prepared by diazotization of the corresponding amine, and that coupling took place on the coupling component with the elimination of a hydrogen ion.

The coupling of benzenediazonium chloride with *m*-phenylenediamine in acid medium gives Basic Orange 2 (Chrysoidine), a dye of limited use on textiles but of importance for coloring paper, leather, and woodstains.

Basic Orange 2

Acid Orange 7 (Orange II) is prepared by coupling diazotized sulfanilic acid with *beta*-naphthol in basic medium. Acid Orange 7 is useful on textile fibers, such as wool, silk, and nylon, as well as on paper and leather.

Acid Orange 7

Other acid monoazo dyes of importance are derived from a variety of sulfonated intermediates.

FD&C Red 40 is made from diazotized cresidine sulfonic acid (4-amino-5-methoxy-2-toluenesulfonic acid) and Schaeffer's acid. Acid Red 337 is made from diazotized *o*-aminobenzotrifluoride and Gamma Acid. The coupling in this case is acidic, and the resulting dye is used for the coloration of nylon.

FD&C Red 40

Acid Red 337

Mordant Black 17 is an example of a type of dye that shows improved fastness on wool when treated on the fiber with a soluble chromium salt. Reaction occurs between the dye and the chromium ion to form a complex of enhanced fastness to washing and light. The component parts of Mordant Black 17

will be recognized as 1-amino-2-naphthol-4-sulfonic acid and *beta*-naphthol.

Mordant Black 17

Another dye of this type is Mordant Brown 4, which is made by diazotizing picramic acid and coupling it in acid medium with toluene-2,4-diamine.

Picramic acid

Mordant Brown 4

Acid Yellow 23 (Tartrazine) contains the pyrazolone nucleus. It is prepared by coupling diazotized sulfanilic acid with 5-oxo-1-(4-sulfophenyl)-2-pyrazoline-3-carboxylic acid or by heating oxidized tartaric acid (1 mole) with phenylhydrazinesulfonic acid (2 moles). Acid Yellow 23 is useful as a dye for basic fibers, paper, and leather. It also is used as a filter dye in photography. The purified dye is used in food coloring as FD&C Yellow #5.

A further example of pyrazolone dye is Acid Red 38. This dye is the chromium complex of the azo compound made by coupling diazotized 2-amino-5-nitrophenol with *meta*-(3-methyl-5-oxo-2-pyrazolin-1-yl)-benzenesulfonamide. It is used as a dye for wool and nylon. Acid Red 38 has one advantage over dyes such as Mordant Black 17 in that the dyer does not have to form the chromium complex by a separate operation during the dyeing process.

Acid Red 38

Disperse Red 1 is a dye for cellulose acetates, nylon, polyesters, and various plastics. It is prepared by coupling diazotized *p*-nitroaniline with *N*-ethyl-*N*-(2-hydroxyethyl)-aniline in acid medium. It has low solubility in water and is supplied in finely divided form mixed with a dispersing agent.

Disperse Red 1

The hue can be deepened into a useful blue by adding electronegative auxochromes in the diazo component, as in Disperse Blue 165.

Pyrazolone

Acid Yellow 23

Disperse Blue 165

Some insoluble azo dyes are used as pigments. Pigment Yellow 3 (Hansa Yellow 10G) illustrates this type. It is prepared by coupling diazotized 4-chloro-2-nitroaniline with *o*-chloroacetoacetanilide.

Pigment Yellow 3

A final example of a monoazo color is an orange fiber-reactive dye, prepared by diazotizing 2-amino-1,5-naphthalenedisulfonic acid, coupling it with *X*-methyl-J-acid, and condensing with cyanuric chloride and then with ammonia. One chlorine atom remains for reaction with textile fibers bearing hydroxyl or amino groups.

Disazo Dyes. Disazo dyes contain two azo linkages. They are formed: (1) by the further diazotization and coupling of monoazo dyes containing free amino groups; (2) by the coupling of two moles of diazonium compound to a coupling component that can couple twice; or (3) by the tetrazotization of a diamine followed by coupling to two moles of coupling component. When two moles of diazonium compound or coupling component are used, they may be the same, or they may be different in structure.

Acid Red 73 is a disazo dye prepared by diazotizing aminoazobenzene and coupling to G Acid.

Acid Red 73

Other disazo dyes of this kind are Acid Red 115 (Cloth Red B), which is made by coupling diazotized aminoazotoluene with R acid, and Acid Black 18. Acid Black 18 will be recognized as having been made by the combination, through diazotization and coupling, of sulfanilic acid, *alpha*-naphthylamine, and H acid.

Cyanuric chloride

Reactive Orange 13

Acid Red 115

Acid Black 18

Solvent Red 27 is a disazo dye used to color gasoline. Its formula indicates that it is made by diazotizing aminoazoxylene and coupling to *beta*-naphthol.

Solvent Red 27

An example of a disazo dye prepared by coupling two different diazonium compounds to a single coupling component is Acid Black 1. This dye is made by coupling diazotized *p*-nitroaniline (acid) and aniline (alkaline) to H acid.

A direct dye of importance is Direct Blue 1 (Sky Blue FF). It is made by coupling tetrazotized *o*-dianisidine twice with Chicago acid (SS acid). Direct Blue 1 is used as a dye for cellulosic textiles, paper, and nylon. It is applied to cellulosic textiles and paper under neutral conditions, and it dyes them by adsorption. It also is used as an acid dye on nylon and must be applied under acidic conditions.

As mentioned above, it is not necessary to couple both diazonium groups of a tetrazotized diamine to the same coupling component. Acid Red 114 is made by coupling tetrazotized *o*-tolidine with G acid and with phenol, followed by formation of an ester of the phenol with *p*-toluenesulfonyl chloride. This dye is used on wool, silk, and leather.

Acid Black 1

o-Dianisidine

Direct Blue 1

o-Tolidine

Acid Red 114

The fastness to light of direct dyes often is improved by the formation of copper chelates. Direct Red 83 is the copper complex of the disazo dye prepared by coupling two moles of diazotized 2-aminophenol-4-sulfonic acid to one mole of J acid urea.

diaminostilbenedisulfonic acid. An example is Direct Yellow 4 (Brilliant Yellow), which is made by coupling tetrazotized diaminostilbenesulfonic acid twice to phenol. Direct Yellow 4 is an important dye for paper; it also is used as an indicator because it turns red in

J acid urea

Direct Red 83

Because dyes of this kind usually are made by reacting the coupling product with a cupric salt in aqueous ammonia, ammonia molecules are shown coordinated with the copper in the complex. This position sometimes may be filled by water or by various amines.

A number of important dyes are based on

strong alkali. Reaction of Direct Yellow 4 with ethyl chloride under pressure gives the ethyl ether, Direct Yellow 4 with ethyl chloride under pressure gives the ethyl ether, Direct Yellow 12 (Chrysophenine), which is an important dye for textiles and paper. Unlike Direct Yellow 4, it is not an indicator.

Diaminostilbenedisulfonic acid

Direct Yellow 4

Direct Yellow 12

Alkylation of phenol couplings is commonly practiced in dye chemistry to eliminate the acid/basic sensitivity of the dye. In addition to *o*-alkylation with ethyl chloride, dimethylsulfate and diethylsulfate are used. The latter is preferred to the toxic dimethylsulfate.

Some stilbene azo dyes are prepared by methods other than diazotization and coupling. Direct Orange 61, for example, is prepared by the condensation in alkaline solution of aminoazobenzene-*m*-sulfonic acid and dinitrostilbenedisulfonic acid. The exact structure of Direct Orange 61 is not known, but the presence of an azo linkage formed by reaction of an amine and a nitro group can be demonstrated.

Aminoazobenzene-*m*-sulfonic acid

Dinitrostilbenedisulfonic acid

Polyazo Dyes. Azo dyes can be made with three, four, or more azo linkages. Those with three (trisazo) and four (tetrakisazo) are quite common. An example of a trisazo dye for cotton is Direct Black 80. It is made by the following sequence: *p*-aminoacetanilide is diazotized and coupled alkaline to Gamma Acid, the acetyl group is removed by hydrolysis, and the resulting diamine is tetrazotized and coupled with 1,6 Cleve's Acid to Gamma Acid (alkaline).

Direct Black 80

Manufacturing Processes for Azo Dyes

The typical azo plant is a three-story design. The makeup tubs for solutions and diazotizations are on the third floor and the coupling tubs on the second floor. The diazotized amine solution flows by gravity from the third floor tub to the second floor tub where coupling takes place. From there, the precipitated dye slurry is transferred to a blowcase on the first floor and blown by air to the filtration equipment. This scheme is diagrammed in Fig. 23.3 and may be illustrated by the processing of Disperse Yellow 3, an important monoazo disperse yellow dye made by coupling diazotized *p*-aminoacetanilide with *p*-cresol.

Disperse Yellow 3

p-Aminoacetanilide is suspended in water in a wooden, brick-lined, or rubber-lined tub (Tub 2) equipped with an agitator. Two to 2.5 moles of hydrochloric acid per mole of amine are added, and the whole is brought to a boil by passing in live steam. As soon as the *p*-aminoacetanilide has dissolved, ice is added. One mole of sodium nitrite, in solution in Tank 1, is run in slowly for about 2 hours. In the meantime, two moles of *p*-cresol are dissolved in an excess of soda ash in water (Tub 3), and to this the contents of Tub 2 are added with stirring. Enough ice is added to keep the temperature at about 5°C. The dye separates in part as the coupling proceeds. When the coupling is complete, the contents of Tub 3 are warmed with steam.

The suspension of dye is run into a plate-and-frame filter press. The filter cake is washed and freed from most of the adhering mother liquor by blowing it with compressed air while still in the press. The moist presscake is discharged into carts and transported to

Fig. 23.3. Schematic diagram for manufacturing a monoazo dye.

another vessel, where it is liquefied with a dispersant. The dye is then dispersed by passing the slurry through a wet-grinding mill until the particle size is about 1 micron. More dispersant or other diluent is added to make it equal in color strength to a predetermined standard, and it is finally spray-dried.

If the dye is in a dye class other than Disperse that does not require dispersion milling to a fine small particle size, the wet presscake is discharged into shallow trays that are placed in a circulating air dryer, wherein the moisture is removed at temperatures between 50°C and 120°C. Vacuum dryers and drum dryers also may be used. The dried dye is ground and mixed with a diluent to "standardize" it. Dilution is necessary because batches differ in their content of pure dye; if it were sold "as is," users would have to adjust their dyeing recipes for each batch. Therefore, uniformity is assured by dilution to a standard strength.

Some typical equipment used in dye manufacture is shown in Figs. 23.4 and 23.5.

(a)

(b)

Fig. 23.4. (a) Filter press room; (b) internals of a plate-and-frame filter. (*Courtesy Crompton and Knowles Corp.*)

Fig. 23.5. Glass-lined steel reactor. (*Courtesy Crompton and Knowles Corp.*)

Triphenylmethane Dyes

Triphenylmethane dyes are characterized by the resonance-stabilized chromophore shown below. They are among the strongest and brightest of synthetic dyes, but usually do not exhibit good fastness to light. An exception is found with acrylic fibers, which are dyed in bright lightfast hues by selected triphenylmethane dyes.

The triphenylmethane chromophore

Basic Red 9 (Pararosaniline) is the simplest triphenylmethane dye. It is prepared by the condensation of one mole of *p*-toluidine and two moles of aniline, usually by heating in nitrobenzene, which serves as solvent and oxidant. The first product is a colorless carbinol, which is converted to the dye with hydrochloric acid.

Colorless carbinol

Basic Red 9

Closely related to Basic Red 9 is Basic Violet 14 (Fuchsine). This dye is made by condensing mol mole each of aniline, *p*-toluidine, and *o*-toluidine in nitrobenzene. The dye is formed on neutralizing the carbinol with hydrochloric acid.

A second method for the synthesis of triphenylmethane dyes is illustrated with Basic Green 4 (Malachite Green). Benzaldehyde (one mole) is condensed with two moles of *N,N*-dimethylaniline to form a leuco base, which is oxidized with lead peroxide and treated with hydrochloric acid to give the dye.

Basic Green 4 is used as a dye for bast fibers, acrylic fibers, leather, paper, and lacquers. When *o*-chlorobenzaldehyde is used in place of benzaldehyde, the product is Basic Blue 1, a more alkali-resistant dye that finds use on tannin-mordanted cotton, wool, leather, and paper.

Basic Violet 14

Basic Green 4

Basic Blue 1

Ketone condensation, a third method of manufacture of triphenylmethane dyes, is illustrated with Basic Violet 3 (Crystal Violet). Michler's Ketone, made by passing phosgene into *N,N*-dimethylaniline, is condensed with *N,N*-dimethylaniline using phosphorus oxychloride. A carbinol is formed, which is converted to the dye with hydrochloric acid.

Michler's Ketone

Basic Violet 3

Not all triphenylmethane dyes are basic. Acid Green 3, for example, is a dye for wool, silk, and leather. There are two sulfonate groups in the molecule, one of which forms an inner salt with the positive charge on the chromophore.

Acid Green 3

Closely related to the triphenylmethane dyes are the diphenylmethane dyes. One of these, Basic Yellow 37, is used for coloring the yellow pages of telephone books.

Basic Yellow 37

Xanthene Dyes

The chromophore of the aminoxanthene dyes is the resonance-stabilized structure shown below, where R is H, or alkyl, or aryl. The hydroxyxanthenes can be stabilized by the loss of a proton, forming an uncharged system in which the chromophore is a quinoid structure.

The aminoxanthene chromophore

The hydroxyanthene chromophore

Acid Yellow 73 (Fluororescein), Acid Red 87 (Eosine), and Basic Violet 10 (Rhodamine B) each represent a series of xanthene dyes.

Acid Yellow 73 is made by heating phthalic anhydride (1 mole) and resorcinol (2 moles) in an iron kettle. The temperature is regulated by hot oil in the kettle jacket and held at 220°C for 7 hours. The melt is dissolved in caustic soda, and the product is precipitated by acidification. It is a yellow-red powder that

fluoresces green-yellow in alkaline solution. The soium salt, called Uranine, is used to trace the underground flow of water.

Phthalic anhydride Resorcinol

Acid Yellow 73

When Acid Yellow 73 is dissolved in alcohol and treated, while warm, with bromine, four equivalents are absorbed to form Acid Red 87 (Eosine). This is used to a limited extent on wool; its primary uses are for paper and inks. Other dyes related to Acid Red 87 are Acid Red 91 (dinitrodibromofluorescein), Acid Red 92 (Phloxine; tetrabromotetrachlorofluorescein), and Acid Red 51 (Erythrosine; tetraiodofluorescein).

Acid Red 87

Basic Violet 10 (Rhodamine B) is a dye for bast fibers, mordanted cotton, leather, and paper. Its salts with certain acids and acid dyes provide pigments (Pigment Violet 1, Pigment Red 173) and solvent dyes (Solvent Red 35, Solvent Red 109). Basic Violet 10 is prepared by condensing phthalic anhydride with *m*-diethylaminophenol and treating with hydrochloric acid. This is the general preparative route for rhodamines.

Basic Violet 10

A variation, Acid Red 388, is an important fluorescent dye that is superior to Uranine for water tracing. Its preparation is similar to the preceding one, with condensation of trimellitic anhydride with *m*-diethylaminophenol.

Trimellitic anhydride

Acid Red 388

Related to the xanthene dyes are the *acridines*, *azines*, *oxazines*, and *thiazines*. Acridine Yellow results from the fusion of toluene-2,4-diamine with glycerin and oxalic acid, followed by oxidation with ferric chloride. Basic Orange 15 (Phosphine) is a by-product of the manufacture of Basic Violet 14 (Fuchsine) and is an acridine dye.

Acridine Yellow

Basic Orange 15

The preparation of azine dyes is illustrated with Safranine B. *p*-Phenylenediamine and aniline are reacted to form indamine. Further reaction of indamine with aniline under oxidizing conditions, followed by treatment with hydrochloric acid, gives Safranine B.

Basic Blue 9 (Methylene Blue) illustrates the thiazine dye class.

Basic Blue 9

Anthraquinone and Related Dyes

Dyes based on anthraquinone and related polycyclic aromatic quinones are of great importance. Many of the most lightfast acid, mordant, disperse, and vat dyes are of this kind. The chromophore is the carbonyl group,

Phenylenediamine Indamine

Safranine B

Indamine (above) is an example of a *quinoneimine* dye. Dyes of the indamine types are used in color photography; the dye is formed during development of the film.

Basic Blue 3 is an oxazine dye that is important for dyeing acrylics greenish-blue shades. It is prepared by condensing nitrosated diethylamino-*m*-anisidine and diethylamino-*m*-phenol under acidic conditions. The methoxyl group is eliminated during the condensation.

$\diagdown$C=O, which may be present once or several times.

Anthraquinone Acid Dyes. Anthraquinone acid dyes are illustrated by Acid Blue 25 and Acid Green 25. These dyes are water-soluble anthraquinone derivatives that are used to dye wool, silk, nylon, leather, and paper.

Basic Blue 3

The starting material for Acid Blue 25 is 1-aminoanthraquinone, which is sulfonated and brominated to give bromamine acid. The copper-catalyzed condensation with aniline then yields Acid Blue 25, which is isolated as the sodium salt.

Bromamine acid

Acid Blue 25

Substitution of aniline by other aromatic amines leads to many other blue acid dyes.

Acid Green 25 (Alizarine Cyanine Green G) is prepared by sulfonation of Solvent Green 3, which in turn is obtained by condensation of quinizarin and *p*-toluidine.

In addition to its use as a dye for wood and nylon, Acid Green 25 is used for the coloration of automobile antifreeze solutions.

Anthraquinone Disperse Dyes. Anthraquinone disperse dyes are illustrated by Disperse Blue 56, Disperse Red 60, and Disperse Red 86. These dyes are bright, lightfast shades on polyester.

Disperse Blue 56

Disperse Red 86

Quinizarin

Solvent Green 3

Acid Green 25

Disperse Red 60

Anthraquinone Vat Dyes. Many of the best vat dyes are derivatives of anthraquinone or related compounds. A relatively simple dye of this class is Vat Yellow 3, which is made by benzoylation of 1,5-diaminoanthraquinone. It is a yellow pigment that is used as such when properly ground and dried. As a vat dye, it usually is supplied as an aqueous paste. Reduction with sodium hydrosulfite in caustic soda solution gives the alkali-soluble anthrahydroquinone, which has affinity for cellulosic fibers. After application to the fiber, the insoluble dye is re-formed by oxidation, as described earlier under "Vat Dyes."

Vat Yellow 3

Perhaps the best known of the anthraquinone vat dyes is Vat Blue 4 (Indanthrone). This attractive dye is made by the fusion of 2-aminoanthraquinone with caustic potash. The fastness to chlorine bleaching of Vat Bblue 4 is improved by chlorination. A number of chlorinated indanthrones are in commercial use and make up the most important group of fast vat blues. The preparative route for Vat Blue 4 is outlined below:

Silver salt

2-Aminoanthraquinone

Vat Blue 4

Related to Vat Blue 4 are Vat Yellow 1 (Flavanthrone) and Vat Orange 9 (Pyranthrone), as well as other important dyes.

Vat Yellow 1

Vat Orange 9

One of the most attractive of all vat dyes is Vat Green 1 (Jade Green). This dye is a derivative of dibenzanthrone, which is itself a vat dye (Vat Blue 20 of Violanthrone).

Vat Blue 20

Vat Green 1

Anthraquinone vat dye manufacture is illustrated by the preparation of Vat Blue 20. Treatment of anthraquinone in 82 percent sulfuric acid with metal powder (iron, copper) and glycerin gives benzanthrone, probably by the series of steps shown below. Condensation of benzanthrone with caustic potash in isopropanol and sodium acetate gives dibenzanthronyl, which undergoes ring closure to dibenzanthrone upon further heating. The dibenzanthrone usually is dissolved in sulfuric acid and reprecipitated as a paste of fine particle size by dilution with water. The paste is washed free of acid and standardized as Vat Blue 20. The equipment for this series of reactions must be able to resist the strongly corrosive conditions and must be externally heated and cooled. Cast iron and stainless steel closed vessels are used (Fig. 23.6).

Phthalic anhydride *o*-Benzoylbenzoic acid

Anthraquinone Anthrone Intermediate

Benzanthrone Dibenzanthronyl Dibenzanthrone Vat Blue 20

Fig. 23.6. Schematic diagram for manufacturing dibenzanthrone.

Indigoid and Thioindigoid Dyes

The parent compound of the indigoid dye class, indigo, has been in use as a vat dyes since ancient times. Natural indigo, which is obtained from a plant species, *Indigofera*, can be reduced by fermentation as well as by the modern method using sodium hydrosulfite and caustic soda. Once reduced, it can be applied to cellulosic fibers and then oxidized by air to produce rich blue dyeings. Natural indigo no longer is important, but large quantities of the synthetic dye are used to produce low-cost blues on cotton. The oxidized form also is used as a pigment. The *Colour Index* designation for indigo is Vat Blue 1.

A modern synthesis of Vat Blue 1 involves first the preparation of phenylglycine. To a water solution of sodium bisulfite are added formaldehyde, aniline, and sodium cyanide. The nitrile of phenylglycine that is formed is isolated and washed free of sulfite. Warming of the nitrile in an alkaline slurry then gives a solution of phenylglycine as its sodium salt. Ammonia is evolved during this step.

In the second step the phenylglycine salt is condensed by heating with a eutectic mixture of potassium and sodium hydroxides in the presence of sodamide, which removes the water formed. The reaction temperature is 200 to 220°C. The product, which is formed in high yield, is indoxyl.

Indoxyl

After the fusion step, water is added, and air is blown into the alkaline solution of indoxyl. Indigo forms and separates from solution. It is filtered, washed, and standardized as a paste of about 20 percent solids.

Nitrile of phenylglycine Phenylglycine

Indoxyl Vat Blue 1

Because very large volumes of indigo are produced, it has been worthwhile to devise continuous processes. Such processes were pioneered by the former I. G. Farbenindustrie in Germany.

Vat Blue 1 is available as a water-soluble salt of the reduced dye. When applied to the fiber and treated with a solution of sodium nitrite or sodium bichromate, it reverts to the insoluble dye.

Solubilized Vat Blue 1

Important derivatives of indigo are the chloro and bromo derivatives. Vat Blue 5, which is tetrabromoindigo, is similar to Vat Blue 1 in shade but is considerably brighter. It is made by bromination of indigo in acetic acid.

Vat Red 41 (Thioindigo) is made by the reaction of thiosalicyclic acid with chloroacetic acid, followed by fusion with caustic soda and oxidation by air. Thiosalicylic acid is made from anthranilic acid by diazotization and reaction with hydrogen sulfide.

Vat Red 41 is a dull bluish red on cotton. On polyester fibers it produces bright pink shades. Many of its derivatives, however, are of greater commercial importance on cotton. An example is Vat Orange 5.

Vat Red 41

Vat Orange 5

Sulfur Dyes

Sulfur dyes are dyes of unknown constitution that can be applied to fibers when reduced with sodium sulfide. Most of them are insoluble in water before reduction, but after reduction they are soluble and can be

Anthranilic acid Thiosalicylic acid

o-(Carboxymethylthio) benzoic acid Thioindoxyl

adsorbed by fibers and then oxidized to an insoluble form with air. Although structures cannot be written for the sulfur dyes, the methods for reproducing individual types are well established.

m-Diamines (toluene-2,4-diamine, for example) fused with sulfur cause the evolution of hydrogen sulfide and produce a melt that yields brown sulfur dyes. Redder shades are obtained by using derivatives of phenazine, such as the compound below, which produces Sulfur Red 6.

Intermediate for Sulfur Red 6

Blue sulfur dyes are obtained by treating diphenylamine derivatives with sodium polysulfide. Sulfur Blue 7, Sulfur Blue 9, and Sulfur Blue 13 are made from the intermediates below:

Intermediate for Sulfur Blue 7

Intermediate for Sulfur Blue 9

Intermediate for Sulfur Blue 13

Sulfur Black 1 is prepared by heating 2,4-dinitrophenol with sodium polysulfide. The fusion mass is dissolved in water and blown with air until all the dye has separated. Then it is filtered, washed, and dried.

Intermediate for Sulfur Black 1

Vat Blue 43 (Hydron Blue R) is a kind of sulfur dye that, unlike typical sulfur dyes, may be reduced by sodium hydrosulfite without destruction. To prepare Vat Blue 43, carbazole is condensed with *p*-nitrosophenol in sulfuric acid to give an indophenol, which is reduced and then fused with polysulfide.

Carbazole *p*-Nitrosophenol

Condensation product

→Vat Blue 43

Reduced condensation product

Vat Blue 42 (Hydron Blue G) gives a greener shade than Vat Blue 43. It is prepared by using *N*-ethylcarbazole in place of carbazole.

Phthalocyanines

The phthalocyanines constitute an important class of synthetic pigments and dyes. The parent compound is Pigment Blue 16 (phthalocyanine). One method of preparation involves the fusion of phthalonitrile with cyclohexylamine in an inert solvent. The two central hydrogen atoms can be replaced by metals such as copper, nickel, iron, and cobalt.

In actual practice, the metal derivatives are not made from the metal-free compound, but are synthesized directly. Pigment Blue 15 (copper phthalocyanine), for example, is made by the fusion of phthalonitrile with copper metal or a copper salt. An alternative method involves the fusion of phthalic anhydride with urea and a copper salt in the presence of a molybdenum catalyst. Pigment Blue 15 and its polychlorinated derivative, Pigment Green 7, are the most important of the phthalocyanines. They are distinguished by great brilliance, strength, and stability, and are used in every field in which colored pigments are used.

Phthalonitrile

Phthalic anhydride

Pigment Blue 15

A number of phthalocyanine derivatives can be made by replacing the hydrogen atoms on the benzene rings. Pigment Green 7, mentioned above, is an example. Vat Blue 29 is a partially sulfonated cobalt phthalocyanine. Reduction with sodium hydrosulfite in caustic soda solution converts it to a soluble form that can be applied to textiles and reoxidized to the pigment. The structure of the reduced form is not known.

Direct Blue 86 is copper phthalocyanine that has been sulfonated to the extent of between two and three sulfonic acid groups per molecule. It produces very bright blue-green shades on cotton, viscose, and paper. Solvent Blue 25 is a spirit-soluble derivative of copper phthalocyanine, useful for inks, lacquers, and stains. It is a reaction product of the tetrasulfonyl chloride of copper phthalocyanine with 4-methylpentylamine.

PRODUCTION AND SALES OF DYES

Table 23.4 gives the production and sales data for synthetic dyes in the United States for the year 1988, the last year such data are available. These figures are taken from a report of the United States Tariff Commission.[7]

The Tariff Commission reported that approximately two-thirds of the dyes consumed in the United States are used by the textile industry, one-sixth to color paper and the rest chiefly in the production of organic pigments and in dyeing leather and plastics.

Imports from Western Europe accounted for 60 percent of the total U.S. imports in 1989. Most of these imports represented intracompany sales between European dye manufacturers and their U.S. subsidiaries. In 1989, the largest single country supplying

TABLE 23.4 1988 U.S. Production and Sales of Dyes by Application Class

		Sales	
	Production (1000 lb)	*Quantity (1000 lb)*	*Value (1000 $)*
Acid	17,779	15,372	75,338
Basic	13,846	12,591	72,575
Direct	41,616	40,053	98,134
Disperse	30,064	21,874	99,689
Fiber-reactive	19,235	18,669	107,589
FD&C	6,973	6,625	51,017
Mordant	56	39	178
Solvent	12,448	8.921	43,012
Vat	30,966	30,355	58,648
Other	30,023	26,526	57,906
Total	203,006	181,025	664,086

imports (in terms of value) was Japan at 39 percent, followed by Germany at 32 percent and Switzerland at 12 percent.[8]

NEW DEVELOPMENT IN DYES

As in the past, dye manufacturers will continue to react to any new development in fiber variations. There continues to be emphasis on the development of dyes with improved fastness properties and ease of application, as well as low-impact, environmentally improved products.

The trend has been to provide dyes in forms that reduce dust exposure, both at the manufacturing sites and in the consumers' plants. More dyes will become available in granular or bead forms. Liquids, too, are becoming increasingly important, particularly where automated liquid metering and dispensing systems are used.

More data will be provided concerning the chemical structures, diluents, and toxicity of individual dyes. The U.S. Toxic Substances Control Act requires that any new substance be registered with the Environmental Protection Agency by the filling of a pre-manufacturing notification (PMN). In most cases, toxicity test data are required before the PMN is granted.

All dyes and intermediates require Material Safety Data Sheets (MSDS) which provide physical, fire and explosion, health hazard, and reactivity data as well as EPA registration numbers, a listing of hazardous ingredients, spill or leak procedures, special protective equipment information, and special precautions to be taken in handling and storage.

In addition, the dye industry has been required to meet new environmental regulations during the past decade. These regulations will become stricter in the future; so the cost of dye manufacture will continue to increase to comply with pollution abatement requirements.

Dyes will be produced with more consistent quality and will vary little from lot to lot, as both manufacturers and users become more involved in quality improvement programs, and instrument techniques make exact color matching possible.

The challenge for ever improving consistency and reproducibility batch to batch will be met by quality improvement programs, more automation in the synthetic process, and more stringent quality control limits.

REFERENCES

1. Society of Dye & Colourists and American Association of Textile Chemists & Colorists, *Colour Index*, 3rd ed., 5 Vol., 1971; 1st Revision of Vol. 5, 1975; Vol. 6 (1st Supplement to Vols. 1–4), 1975; Vol. 7 (Supplement to Vols. 1–4 and 6), 1982; Vol. 8 (Supplement to Vols. 1–4, 6, and 7), 1987.
2. U.S. patents, 3,709.902; 3,784,599; 3,935,182; 3,996,282; 4,065,500.
3. Mecenier, P. L., *Textile Chemist & Colorist*, **2**, 386 (1970).
4. Garondy, E. J., *Textile Chemist & Colorist*, **10**, 105 (1978).
5. Rast, G. E., and Steiner, R. I., U.S. Patent 3,578,652, 1971.
6. "Imports of Benzenoid Chemicals and Products 1983," *USITC Publication 1548* (July 1984).
7. "Synthetic Organic Chemicals," *USITC Publication 2219* (Sept. 1989).
8. Rose, C. J., Oppenheimer & Co., Inc., Report No. 90-462 (Apr. 5, 1990).
9. Tutner, R., *Textile Chemist & Colorist*, **19** (8), 16 (1987).
10. Turner, R., *Textile Chemist & Colorist*, **19** (9), 69 (1987).
11. Turner, R., *Textile Chemist & Colorist*, **20** (10), 25 (1988).
12. Turner, R., *Textile Chemist & Colorist*, **21** (8), 23 (1989).
13. *Dyeing Primer*, **12** (Nos. 1–12) (1980).
14. Kauffman, G., *Industrial Chemicals*, p. 26 (1988).
15. Fierz-David, H. E., and Blangey, L., *Fundamental Processes of Dye Chemistry*, trans. by P. W. Vittum, John Wiley & Sons, New York, 1949.
16. Groggins, P. H. (Ed.), *Unit Processes in Organic Synthesis*, 5th ed., McGraw-Hill, New York, 1958.
17. I. G. Farbenindustrie, FIAT 764 (PB 60946), "Dyestuffs Manufacturing Processes of I. G. Farbinindustrie A. G." (1947).

18. Lubs, H. A. (Ed.), *The Chemistry of Synthetic Dyes & Pigments*, Reinhold, New York, 1955; Hafner reprint, New York, 1965.
19. Peters, R. H., *Textile Chemistry*, Vol. III, *The Physical Chemistry of Dyeing*, Elsevier, New York, 1975.
20. Zollinger, H., *Azo and Diazo Chemistry*, Interscience Publishers, New York, 1961.
21. Venkataraman, K. (Ed.), *The Chemistry of Synthetic Dyes*, 8 Vol., Academic Press, New York, 1952–78.
22. Venkataraman, K. (Ed.), *The Analytical Chemistry of Synthetic Dyes*, Wiley-Interscience, New York, 1977.
23. Hollen, N., and Saddler, J., *Textiles*, 3rd ed., Macmillan, New York, 1968.
24. Trotman, E. R., *Dyeing and Chemical Technology of Textile Fibers*, 3rd ed., Griffin, London, 1964.
25. Mark, H. F., Atlas, S. M., and Cernia, E. (Eds.), *Man Made Fibers: Science and Technology*, 3 Vols., Wiley-Interscience, New York, 1967–68.
26. Robertson, S. M., *Dyes from Plants*, Van Nostrand Reinhold, New York, 1973.
27. Wilcock, J., and Ashworth, R. (Eds.), *Whittaker's: Dyeing with Coal Tar Dyestuffs*, Van Nostrand, New York, 1969.

24

Industrial Fermentation: Principles, Processes, and Products

Arthur E. Humphrey* and S. Edward Lee**

EVOLUTION OF MODERN FERMENTATION PROCESSES

Humans were well aware of fermentations, even though they had little knowledge of what caused them, long before they were able to record such an awareness. Sometime in prerecorded history, people discovered that meat allowed to stand a few days was more pleasing to the taste than meat eaten soon after the kill. They also were aware that intoxicating drinks could be made from grains and fruits. The aging of meat and the manufacture of alcoholic beverages were human's first uses of fermentation.

Without even knowing that microorganisms existed, ancient people learned to put them to work. The anciet art of cheese-making involves fermentation of milk. For thousands of years, the soy sauces of China and Japan have been made from fermented beans. For centuries, the Balkan people have enjoyed fermented milk, or yogurt, and members of Central Asian tribes have found equal pleasure in sour camel's milk, or kumiss. Bread, which has been known almost as long as agriculture itself, involves a yeast fermentation. Loaves of bread have been found in Egyptian pyramids built 6000 years ago.

The discovery of fruit fermentation was made so long ago that the ancient Greeks believed wine had been invented by the god Dionysus. The manufacture of beer is only slightly less ancient that that of wine. A Mesopotamian clay tablet written in Sumerian-Akkadian about 500 years before Christ tells us that brewing was a well-established profession 1500 years earlier. An Assyrian tablet of 2000 B.C. lists beer among the commodities that Noah took aboard his ark. Egyptian documents dating back to the Fourth Dynasty, about 2500 B.C., describe the malting of barley and the fermentation of beer. Kui, a Chinese rice beer, has been traced to 2300 B.C.

During the Middle Ages, experimenters learned how to improve the taste of wine, bread, beer, and cheese. Yet, after thousands of years of experience people still did not realize that in fermentations they were dealing with microorganisms. It was not until 1857 that Pasteur proved alcoholic fermentation was brought about by yeasts, and that yeasts

*Lehigh University.
**Pfizer Central Research.

were living cells. The discovery marked a turning point in medical history and the birth of microbiology.

The inheritors of Pasteur's knowledge sought to use microbes as production workers in industry. The production of bakers' yeast in deep, aerated tanks was developed toward the end of the nineteenth century. During World War I, Chaim Weizmann used a bacterial cousin of the gas gangrene microbe to convert maize mash into acetone, which is essential in the manufacture of the explosive cordite. In 1923, Pfizer opened the world's first successfull plant for citric acid fermentation. The process involved a fermentation utilizing the mold *Aspergillus niger* whereby ordinary sugar was transformed into citric acid.

Other industrial chemicals produced by fermentation were found subsequently, and the processes were reduced to commercial practice. These processes included butanol, acetic acid, oxalic acid, gluconic acid, fumaric acid, and many more.

In 1928 Alexander Fleming, working with *Staphylococcus aureus*, a bacterium that causes boils, observed a strange fact. A mold of the *Penicillium* family grew as a contaminant on a Petri dish inoculated with *Staphylococcus aureus*; a clear zone was observed where the *Staphylococcus* organism in the vicinity of the contaminating mold has been killed. Fleming nurtured the mold and then extracted a chemical from it that killed the bacteria. He named the extracted material penicillin.

Fleming's discovery received little notice until two Oxford University experimenters, under the stress of World War II, resolved to find an antibacterial agent of wider activity than the sulfa drugs. These two British workers, Dr. Howard Florey and Dr. Ernst Chain, set out to produce a therapeutic agent capable of saving the lives of war casualties. Their first candidate was the *Penicillium notatum* mold preserved from Fleming's studies. Penicillin turned out to be exactly what they were looking for. The American pharmaceutical industry helped them solve their difficulties in mass-producing the antibiotic. Three American companies led the way—Merck, Pfizer, and Squibb.

Initially, *Penicillium* was surface-cultured in flasks. A chance discovery in a Peoria market provided the major breakthrough. There, a government worker found a moldy cantaloupe on which was growing a new strain of penicillium, *Penicillium chrysogenum*, which would thrive when cultured in deep, aerated tanks and which gave 200 times more penicillin than did Fleming's mold.

Other antibiotics were quick to appear. Streptomycin was next, an antibiotic that was particularly effective against the causative organism of tuberculosis. The search was now on. Antibiotic prospectors combed the earth for organisms that produced different and more useful antibiotics. The list of these antibiotics is long today and includes such important antibiotics as chloramphenicol, the tetracyclines, bacitracin, erythromycin, novobiocin, nystatin, kanamycin, and many others.

Future Developments

Progress in fermentation is continuing at an ever-increasing pace. Each year, new products are added to the list of compounds derived from fermentation. Several vitamins now are produced routinely, employing fermentation steps in their synthesis. Outstanding examples are B_2 (riboflavin), B_{12} (cyanocobalamin), and C (ascorbic acid). Some of the more interesting fermentation processes are the specific dehydrogenations and hydroxylations of the steroid nucleus. These chemical transformations are economical shortcuts used in the manufacture of the anti-arthritic cortisone and its derivatives. Fermentative syntheses of the amino acids L-lysine and L-glutamic acid also are being carried out commercially. Important agricultural uses are found for the new fermentation product gibberellin, a plant-growth regulator; and crystalline inclusions for a species of *Bacillus* are being used as specific insecticides in another agricultural application. A macrolide, avermectin, produced by a microorganism found near a golf course in Japan turned out to be a most potent and important antiparasitic agent. Microbial hydrolysis of cellulose and subsequent fermentation of the resultant sugar to alcohol is

a promising source of liquid fuel. Research is in progress on chemical transformations utilizing fermentation techniques, new fermentative biosyntheses, continuous algal culture, and submerged mammalian-tissue culture. The advent of recombinant DNA technology since the late 1970s has added to the list of fermentation-derived products such complex molecules as insulin, human growth hormone, interferons, hepatitis vaccine, tissue plasminogen activator, erythropoietin, and colony stimulating factors.

FERMENTATION FUNDAMENTALS

Introduction

The engineer's contribution to the development of the penicillin fermentation was a very important one. The outgrowth of the undertaking was the pure-culture technique, carried out in aerated and agitated deep-tank fermentors. This technique, similar to its antecedent used for yeast propagation, introduced to the biochemical process industry refined fermentation equipment capable of being maintained under aseptic conditions even when vigorously aerated. The technique has now been applied widely with minor modifications to the production of other antibiotics, amino acids, steroids, enzymes, and therapeutic proteins.

Of the operations auxiliary to those in the fermentor, engineers have made a major contribution in establishing the theoretical bases for the design of equipment to provide large volumes of sterile medium and air. Close cooperation between biologists and engineers is needed to devise logical methods for screening large number of strains, and translating the the results of shake-flask and pilot-plant experiments to production vessels.

Fermentation Unit Operations

Fermentation processes have in common many of the familiar chemical engineering unit operations. For example, aerobic fermentations involve the "mixing" of three heterogeneous phases—microorganisms, medium, and air. Other unit operations include "mass transfer" of oxygen from the air to the organisms and "heat transfer" from the fermentation medium.

Analysis of fermentations by the unit-operation technique has added greatly to the understanding of their behavior. This understanding, however, is far from complete. The scale-up of fermentations for instance, is still rather empirical although the sensitive oxygen probes and sensitive gas analysis techniques now available have enabled a more rational approach to scaling-up aerated, non-newtonian fermentations.

Fermentation Unit Processes

Analysis of the many industrial fermentation processes shows that they are common reactions from a chemical as well as a physical viewpoint. Fermentation processes can be classified by the reaction mechanims involved in converting the raw materials into products; these include reductions, simple and complex oxidations, substrate conversions, transformations, hydrolyses, polymerizations, complex biosyntheses, and the formation of cells.

Unit-process classification provides a ready catalogue of the chemical activities and abilities of microorganisms for the biochemist. More important, it offers a logical approach to an examination of fermentation reaction mechanisms.

Microorganisms

Microorganisms are chemically similar to higher plant and animal cells; they perform many of the same biochemical reactions. Generally microorganisms exist as single cells, and they have much simpler nutrient requirements than the higher life forms. Their requirements for growth usually are limited to air, carbon dioxide, and inorganic salts or inorganic salts supplemented by simple sugars. They occur in four main groups:

1. Bacteria
2. Viruses*

*Not a microbial cell by strict definition.

3. Fungi, including yeast and actinomyces
4. Protozoa and algae

Typically, *bacteria* are single cells, either cocci, rods, or spirals, which are 0.5 to 20 μm in size and capable of independent growth. They multiply by simple division. A typical bacterial cell is illustrated in Fig. 24.1.

Viruses are the smallest microbes. They are obligate intracellular parasites of animals, plants, insects, fungi, algae, or bacteria. Bacterial viruses are called bacteriophage. Growth and multiplication take place intracellularly when the virus invades the host cell, taking over the genetic mechanism to direct the host cell to synthesize new viruses. The viruses consist primarily of genetic material, either ribonucleic acid (RNA) or deoxyribonucleic acid (DNA) and a protein coating. Because of the purity of the protein coat, it is possible to crystallize viruses and obtain X-ray diffraction patterns.

The *fungi* are widely spread in nature in environments of lower relative humidity than those that favor bacteria. The metabolism of the fungi is essentially aerobic. They generally form long filamentous nucleated cells 4 to 20 μm wide that are highly branched. Many species have complicated life cycles during which both sexual and asexual spores are formed. These spores usually are quite resistant to environmental changes.

Generally, fungi are free-living saprophytes, but a few are parasites to animal cells, and many are serious pathogens of plants. They have very wide degradative and synthetic capabilities and have proved a fruitful source of industrially important organic acids, antibiotics, and enzymes.

Among the fungi are yeasts, which are elliptical cells 3 to 5 μm by 8 to 15 μm. They differ from most fungi in that growth is by budding. Growth of yeast on solid medium resembles a bacterial colony. Yeasts are industrially important in making alcohol and bakers' yeast.

Actinomycetes are a group of organisms intermediate in properties between bacteria

Fig. 24.1. Detailed structure of a bacterial cell.

and true fungi. Industrially, the group is extremely important as a source of powerful antibiotics for the control of microbial infections of humans, animals, and plants. Their growth in fermentation tanks is very similar to that of fungal fermentations; for this reason, they have been grouped with the fungi in this discussion.

The *protozoa* are widely distributed in fresh and salt water, and found in lesser frequency in soil and animals. They may be unicellular or multicellular and exhibit a wide variety of morphological forms. Usually, they are divided in two main groups: the *algae*, which are all capable of photosynthesis and contain chlorophyll; and the protozoa, which are not photosynthetic and resemble primitive animal cells.

The chemical composition of microorganisms can be quite varied depending upon such factors as the composition of the growth medium, the age of the culture, and the cell growth rate. Table 24.1 lists some of the compositions of different microbial forms.[1]

Mammalian Cells

Mammalian cells have been used for a long time in the manufacture of vaccines for rubella, adenovirus (types 4 and 7), rabies, poliomyelitis, and FMDV (foot-and-mouth disease virus). They have taken on greater importance in industrial fermentation since the advent of recombinant DNA technology. The commercial production of macromolecular proteins, particularly those requiring proper glycosylation and correct folding (e.g., tissue plasminogen activator and erythroprotein) for their biological activities, has accelerated the technical development and understanding of mammalian cell culture. Mammalian cells are quite different from their microbial counterparts, whose general characteristics were described earlier. Mammalian cells, typically 10 to 100 microns in size, are much larger than microbes. They grow much more slowly, doubling in 12 to 60 hours versus 0.5 to 2 hours for the microbes, and they usually require a complex and rich

TABLE 24.1 Chemical Analyses. Dry Weights, and the Populations of Different Microorganisms Obtained in Culture*

Organism	*Composition (% dry weight)* *Protein*	*Nucleic Acid*	*Lipid*	*Population in Culture (numbers/ml)*	*Dry Weight of This Culture (g/100 ml)*	*Comments*
Viruses	50–90	5–50	<1	10^8–10^9	0.0005[a]	Viruses with a lipoprotein sheath may contain 25% lipid
Bacteria	40–70	13–34	10–15	2×10^8–2×10^{11}	0.02–2.9	*Mycobacterium* may contain 30% lipid
Filamentous fungi	10–25	1–3	2–7		3–5	Some *Aspergillus* and *Penicillium* sp. contain 50% lipid
Yeast	40–50	4–10	1–6	1–4×10^8	1–5	Some *Rhodotorula* and *Candida* sp. contain 50% lipid
Small unicellular algae	10–60 (50)	1–5 (3)	4–80 (10)	4–8×10^7	0.4–0.9	Figure ( ) is a commonly found value but the composition varies with the growth conditions

*A. Aiba, A. E. Humphrey and N. Millis, *Biochemical Engineering*, 2nd ed., p. 26. University of Tokyo Press, Tokyo (1973).

[a]For a virus of 200 μm diameter.

medium, often containing some sort of serum component, for growth. This makes the maintenance of asepsis extremely difficult, but most critical, in culturing mammalian cells. Mammalian cells differ from the microbes in another very important manner, in that they do not have a rigid cell wall to protect them against shear and osmotic pressure. This represents an additional challenge to engineers when scaling up a mammalian cell for commercial production. Some mammalian cells require a suitable surface for attachment and growth. Methods and bioreactors that have been developed in the last decade for the growth of the anchorage-dependent cells represent an important technical achievement by the cell culture practioners.

Microbial Activity

All organisms contain the genetic information to produce a wide variety of enzymes and hence produce a great number of chemicals. However, only some enzymes are produced at all times, whereas others are greatly influenced by the substrate. Certain compounds interact with the substrate to repress the translation of genetic information for synthesis. (This process is called repression.) Also, the substrate sometimes reacts with a compound that is a genetic-mechanism repressor and removes its action. (This process is referred to as derepression.) This control is depicted in Fig. 24.2.

These processes allow cells to regulate their enzyme content in direct response to the environment. They prevent the formation of excess end product and superfluous enzymes.

Mutants can be genetically engineered that lack these controls. Such mutants have been changed so that the genetic mechanism no longer is sensitive to a particular controlling metabolite. For industrial processes, such strains with faulty regulation, altered permeability, or metabolic deficiencies may be used to accumulate products.

Strain Development

Although in a normal bacterial population about one mutant arises in 10^6 cells, the rate of mutation is far higher than this because many mutations are lethal and so are not detected, and others involve base changes that do not change the amino acid read from the m-RNA codon. This low rate of spontaneous mutation is unsatisfactory when one is attempting strain improvement. However, mutagens are available that markedly increase the rate of mutation. Recently, tremendous strides have been made in genetic engineering, which allow the creation of mutants through such techniques as:

1. Transduction
2. Transformation
3. Protoplast fusion

We now know that exchange of genetic information between cells is much more common than originally thought. The trick is not only to inject genetic material with the desired

Fig. 24.2. System for the control of enzyme formation in biosynthesis.

characteristics into a cell, but to ensure that it is "expressed" and that an environment is maintained that permits these "new species" to dominate.

Although it is relatively easy to produce mutants, strain improvement requires painstaking effort, ingenuity in devising screening tests, and luck. The difficult problem is to select desirable mutants, and this is accentuated when the parent strain is already a high producer. In this situation, any improvement is likely to be minor and quantitative rather than major and qualitative; major changes are most useful in the very first stages of a development program and in basic research.

Improvement in product formation can be achieved in a number of ways. Mutants can be selected that have the following properties:

1. Defective enzymes in a biosynthetic pathway so that useful intermediates accumulate.
2. Lack of feedback control by the end product itself or by a key intermediate on the pathway to the end product.
3. Defective cell membranes so that intracellular metabolites are excreted rapidly.
4. Lack of some undesirable by-product, such as a pigment.
5. Reduced toxicity to a precursor molecule such as phenyl acetate for penicillin production.

Stock Cultures

In all the industrially important bacteria and fungi, except some yeasts, the main vegetative phase of growth is haploid, so that variation in genetic material arises only from mutations, provided transduction does not occur in bacteria, and provided sexual spores are not formed in fungi. In maintaining stock cultures, genetic change must be minimized; this is best achieved by preventing nuclear divisions because most mutations occur as errors in DNA replication.

The method of choice is to store cells or spores (if these are produced) in sealed ampoules at very low temperatures ($-200°C$) in liquid nitrogen. This method has the great advantage that the culture can be stored almost indefinitely, thawed, and used immediately as an inoculum without loss of viability or diminution in metabolic rate. Cultures kept at $-20°C$ to $-60°C$ are satisfactory but less active than those kept in liquid nitrogen. Although storage at 0°C to 4°C allows some growth, this is better than storage at room temperature. Lyophilization (freeze-drying) is widely used and is very convenient because freeze-dried cultures retain viability without genetic change for years when stored at room temperature. It may be noted that all of these methods are, in effect, techniques to immobilize intracellular water and yet retain viability.

Mammalian Stock Cultures

Stock culture preparation of mammalian cells in general follows that for microbial cells. However, because of the fragility of mammalian cells, extreme caution is recommended in freezing and thawing their stocks. Typically, a controlled, slow freezing with the temperature dropping one-third of a degree every minute us suggested for freezing mammalian cells. Quality control of mammalian cell stocks is much more involved than that for microbial stocks, particularly when the mammalian products are used as injectable human therapeutics. The mammalian cell stocks need to be examined for being free of contamination by bacteria, fungi, and mycoplasma. The culture's identity needs to be tested by banding cytogenetics, isozyme analysis, and product formation. Its safety needs to be examined by various MAP (mouse antibody production), cell line, and animal tests to ensure the absence of active viruses. In addition, the presence or retrovirus should be evaluated by electron microscopy, reverse transcriptase assays, XC and S+L− plaque assays, and chemical induction methods. Its tumorogenecity also needs to be assessed.

Microbial Kinetics

The kinetics of microbial systems may be expressed at four different system levels:

1. Molecular or enzyme level
2. Macromolecular or cellular component level

3. Cellular level
4. Population level

Each level of expression has a unique characteristic that leads to a rather specific kinetic treatment. For example, biological reactions at the molecular level invariably involve enzyme-catalyzed reactions. These reactions, when they occur in solution, behave in a manner similar to that of homogeneous catalyzed chemical reactions. However, enzymes can be attached to inert solid supports or contained within a solid cell structure. In this case, the kinetics are similar to those for heterogeneous catalyzed chemical reactions.

Enzyme Kinetics. In their simplest form, enzyme-catalyzed reactions, occurring in a well-mixed solution, are characterized by the well-known Michaelis-Menten kinetic expression. This relationship depicts the substrate, S, combining reversibly with the enzyme, E, to form an enzyme–substrate complex, ES, that can irreversibly decompose to the product and the enzyme, i.e.:

$$E + S \underset{k_{-1}}{\overset{k_{+1}}{\rightleftarrows}} ES \xrightarrow{k_{+2}} E + P \qquad (24.1)$$

This leads to a kinetic expression for the velocity of the reaction, v, of the following form:

$$v = \frac{v_{\max} S}{K_M + S} = \frac{k_{+2} E_0 S}{\dfrac{k_{-1} + K_{+2}}{k_{+1}} + S} \qquad (24.2)$$

where $v_{\max} = k_{+2} E_0$ is the maximum observable reaction rate at high substrate concentration and, hence, is only limited by the initial enzyme concentration, E_0. K_M is the dissociation constant when

$$k_{+2} \ll k_{-1}, \text{ then } K_M = K_s \qquad (24.3)$$

and

$$ES \xrightarrow{k_2} E + P \text{ is limiting} \qquad (24.4)$$

The saturation constant, $K_s = k_{-1}/k_{+1}$, is an indication of the affinity of the enzyme active site for the substrate.

Basically two kinds of catalytic poisoning or inhibition are considered: inhibition by competition for the active site by a nonreactive substrate, and inhibition by a substance that modifies the enzyme activity but does not compete for the active site. The competitive inhibitor adversely affects the binding of substrate and enzyme, and thus has an effect in increasing K_M. The noncompetitive inhibitor, on the other hand, only affects k_{+2} and $v_{\max}$.

Cell Growth Kinetics. Enzyme kinetic concepts have been utilized to express the kinetic behavior of cell growth on a single limiting substrate. Monod postulated that the growth of cells by binary fission on a single limiting substrate probably had a single limiting reaction step and therefore behaved in a manner analogous to the Michaelis-Menten enzyme kinetics, i.e.:

$$\frac{dX}{dt} = \mu_{\max} X \frac{S}{K_s + S} \qquad (24.5)$$

where X = the cell concentration, S = the growth limiting substrate concentration, t = time, and $\mu_{\max}$ = the maximum growth rate. Since the growth rate of cells, increasing by binary fission, is defined by:

$$\mu = \frac{1}{X}\frac{dX}{dt} \qquad (24.6)$$

The growth kinetics can be expressed as:

$$\mu = \mu_{\max}\left(\frac{S}{K_s + S}\right) \qquad (24.7)$$

For a complete theory of cell growth and substrate utilization, it is necessary to know the relationship between the growth of cells and the utilization of substrate.

This relationship is expressed as a yield, Y, defined as:

$$Y \equiv \frac{dX}{dS} \qquad (24.8)$$

The simplest assumption is that Y is constant. That this is essentially true only at high growth rates will be shown later. From the kinetic equation and the yield expression one obtains the following relationship for substrate

utilization:

$$\frac{dS}{dt} = \frac{1}{Y_G}\frac{dX}{dt} + mX \qquad (24.9)$$

or:

$$\frac{1}{X}\frac{dS}{dt} = \frac{1}{Y_G}\mu + m \qquad (24.10)$$

where m is the maintenance requirement for substrate per unit of cell biomass per unit of time, and Y_G is a true yield constant representing the substrate utilized only for growth.

Recently it has been suggested that the rate of cell increase should be expressed as a net growth rate which involves both growth, μ, and death, δ, i.e.:

$$\frac{dx}{dt} = \mu X - \delta X \qquad (24.11)$$

where:

$$\mu = \mu_{\max}\left(\frac{S}{K_s + S}\right) \qquad (24.12)$$

and:

$$\delta = \delta_{\max}\left(1 - \frac{S}{K'_s + S}\right) \qquad (24.13)$$

Here, cell death is depicted as having a first-order kinetic behavior and as maximal, i.e., $\delta_{\max}$, when the growth limiting substrate is zero, and minimal when the substrate is in excess.

Over the years, numerous models for depicting cell growth have evolved. Several will be discussed here. One such model is that for growth under multiple substrate limitation. It can be expressed as:

$$\mu = \mu_{\max}\left(\frac{S_1}{K_{s1} + S_1}\right)\left(\frac{S_2}{K_{s2} + S_2}\right)\cdots. \qquad (24.14)$$

Multiple substrate limitation frequently occurs in batch growth systems.

Another model for cell growth is that for situations of extremely low substrate concentrations, i.e., $S < K_s$. In this case, the growth kinetic equation reduced to:

$$\mu = KS \qquad (24.15)$$

where K is constant and approximately equal to $\mu_{\max}/K_s$. This kinetic behavior is frequently observed in large single reactor waste treatment systems.

Another situation that commonly occurs in waste treatment is growth under conditions of "shock loading," i.e., conditions in which the substrate rises to a level in which it becomes inhibitory to growth. This situation is frequently modeled by waste treatment designers by the following kind of kinetic expression:

$$\mu = \frac{\mu_{\max}}{1 + \dfrac{K_s}{S} + \dfrac{S}{K_1}} \qquad (24.16)$$

The three basic growth relationships are depicted in Fig. 24.3. In passing, it is interesting to note that the observed $\mu_{\max}$ under conditions of substrate inhibition is a fraction of the true $\mu_{\max}$ if there were no inhibition, i.e.:

$$\frac{\mu \text{ at } d\mu/dS = 0}{\mu_{\max}} = \frac{1}{1 + 2\sqrt{K_s/K_1}} \qquad (24.17)$$

Another common kinetic expression for growth is one that takes into account a substrate diffusional limitation. It has been observed in various waste treatment systems that the K_s for "floc" or sludge growth is greater than that for single cell systems. Further, the growth rate of flocs or sludges appears to be a function of their size. This situation is depicted in Fig. 24.4. The growth

Fig. 24.3. Kinetic expressions for growth.

Fig. 24.4. Effect of sludge size on growth rate.

Fig. 24.5. Effect of temperature on growth.

kinetics in this case can be expressed as:

$$\mu = \frac{\mu_{\max} S}{K_s^{\text{app}} + S} \tag{24.18}$$

and:

$$K_s^{\text{app}} = K_s + K_D \tag{24.19}$$

where $K_D = f$(sludge or cell size, membrane permeability, diffusion coefficient, etc.). When $K_D = 0$, there is no diffusional limitation to substrate uptake.

Temperature Effects. Absolute reaction rate theory has been found applicable to both cell growth and death, i.e.:

$$\frac{dnk}{dT} = \frac{E_a}{RT^2} \tag{24.20}$$

or:

$$dnk = -(E_a/R)d(1/T) \tag{24.21}$$

where k is the rate constant for growth, i.e., $\mu_{\max}$ or death, i.e., $\delta_{\max}$, T is the absolute temperature, E_a is the activation energy for the process, and R is the gas law constant. Generally speaking, E_a for growth, E_G, is the order of 8 to 12,000 cal/g-mole, °K and E_a for death, E_D, is the order of 50 to 100,000 cal/g-mole, °K. This behavior leads to an optimal temperature for growth for a given cell species. This is depicted in Fig. 24.5.

pH Effects. The model for pH effect on growth has been based on the behavior of enzymes. Because enzymes are composed of amino acids, they exhibit "zwitterion" behavior, i.e., they have an acid, base, and neutral form. Cells have been depicted to have an active or neutral form and an inactive base or acid form, i.e.:

$$X^+ \underset{K_1}{\rightleftharpoons} X \underset{K_2}{\rightleftharpoons} X^- \tag{24.22}$$

where K_1 and K_2 are the equilibrium constants of the above reaction. If one defines the growing or active cell fractions as y, then:

$$y \equiv \text{active cell fraction} = \frac{X}{X_0} \tag{24.23}$$

where:

$$X_0 = X^+ + X + X^- \tag{24.24}$$

By analogy to enzyme behavior:

$$y = \frac{1}{1 + \dfrac{[H^+]}{K_1} + \dfrac{K_2}{[H^+]}} \tag{24.25}$$

where $[H^+]$ is hydrogen ion concentration. Figure 24.6 depicts the effect of pH on the active cell fraction.

Fig. 24.6. Effect of pH on growth.

For most cell systems, $pH_{opt} = 6.5 \pm 1$ and $pK_2 - pK_1 = 2 \pm 1$.

Fermentor Configurations

Numerous fermentor configurations may be found in microbial processes. These include:

Batch
Batch-fed
Repeated draw-off
Continuous
- Single stage
- Multiple stage

Continuous with recycle
Continuous with step feeding

For the most part, the antibiotic industry uses batch-type processes. The reason for this stems from the fact that most efficient antibiotic-producing organisms are highly mutated and are readily replaced by fast-growing, less efficient antibiotic producers in a continuous culture. In order to avoid substrate repression or inhibition, some batch processes are continuously fed concentrated substrate on demand during the course of the batch cycle. This is referred to as a "batch-fed" fermentation. The production of bakers' yeast is an example of a batch-fed process. In some highly mycelial antibiotic fermentation, 20 to 40 percent draw-off followed by fresh media makeup is practiced. In the trade, this is referred to as a "repeated draw-off" process. Strict continuous processes are practiced only in processes for the production of biomass for feed or food and the treatment of wastes.

F = flow rate, liters/hr
V = volume, liters
D ≡ F/V = dilution rate, hr^{-1}
X = cell conc., g/liter
S = substrate conc., g/liter

Fig. 24.7. Typical single-stage continuous fermentor.

Contiuous biomass-producing systems are usually single-stage reactors without recycle. Waste treatment systems usually use multistage systems with recycle. They frequently use step feeding of several stages to improve system stability.

A typical single-stage continuous fermentor is depicted in Fig. 24.7.

A biomass material balance around the reactor yields the following relationship. At steady state:

$$\frac{dX}{dt} = 0 \text{ (steady state)} = \frac{F}{V}(0 - X) + \mu X \tag{24.26}$$

from which:

$$\mu = \frac{F}{V} \equiv D \tag{24.27}$$

This means that the dilution rate or the nominal residence time of the fermentor sets the growth rate of the biomass in the fermentor. A change in the dilution rate causes a change in the growth rate.

A similar balance around the fermentor at steady state for the growth limiting substrate, assuming the overall yield, Y, is constant, gives:

$$\frac{dS}{dt} = 0 \text{(steady state)} = \frac{F}{V}(S_0 - S) - \mu\frac{X}{Y} \tag{24.28}$$

from which:

$$(S_0 - S) = \frac{X}{Y} \tag{24.29}$$

Substituting D for μ in Monod growth kinetics yields:

$$S = \frac{K_s D}{\mu_{max} - D} \tag{24.30}$$

and substituting $S_0 - X/Y$ for S gives:

$$X = Y\left(S_0 - \frac{K_s D}{\mu_{max} - D}\right) \tag{24.31}$$

Strictly speaking, Y is not a constant, particularly at conditions of low growth rate because

TABLE 24.2 Effect of Maintenance on Overall Cell Yield in Continuous Culture

	Cell Yield, Y, g/g		
D, hr^{-1}	$m = 0.01$	$m = 0.02$	$m = 0.05$
0.01	0.333	0.250	0.143
0.02	0.400	0.333	0.222
0.05	0.455	0.416	0.333
0.10	0.476	0.455	0.400
0.20	0.488	0.476	0.444
0.50	0.495	0.490	0.476

the maintenance requirement, m, of the biomass utilizes substrate without producing biomass. The steady-state substrate balance equation for a single stage fermenter can be restated as:

$$D(S_0 - S) = \frac{\mu X}{Y_G} + mX \qquad (24.32)$$

Table 24.2 illustrates the effect of maintenance and cell growth or reactor dilution rate on the overall cell yield.

Because most single-stage continuous fermentors are used to produce biomass, they usually are operated to optimize the biomass productivity. The unit volume biomass productivity of such a reactor is defined as DX. This unit volume productivity can be expressed as

$$DX = DY\left(S_0 - \frac{K_s D}{\mu_{max} - D}\right) \qquad (24.33)$$

By taking the first derivative of the productivity expression with respect to the dilution rate and setting it equal to zero, the dilution rate of maximum productivity, D_m, can be found as:

$$D_m = \mu_{max}[1 - \sqrt{K_s/(K_s + S_0)}] \qquad (24.34)$$

and the maximum productivity, $D_m X$, as:

$$D_m X = \mu_{max}[1 - \sqrt{K_s/(K_s + S_0)}] \cdot Y[S_0 - K_s D/(\mu_{max} - D)] \qquad (24.35)$$

The behavior of these relationships is depicted in Fig. 24.8. Note that for this system "wash out" of the biomass occurs when D approaches the maximum growth rate, i.e., $D \rightarrow \mu_{max}$.

Fig. 24.8. Relationships for a continuous single-stage fermentor.

Biomass recycle frequently is used in fermentors as a way of increasing the unit productivity. In order for the system to operate successfully, a biomass concentrator must be connected to the unit (see Fig. 24.9.).

Usually a centrifuge or settling tank is used as the concentrator. The increased productivity achieved with a recycle fermentor depends upon the recycle ration, r, and the cell concentration factor, $C = X_r/X$, achieved in the concentrator. Equations expressing the recycle system behavior are derived from material balances around the reactor, i.e., for the cell biomass balance at steady state:

$$\frac{dX}{dt} = 0 = \frac{F}{V}(0) + \frac{rF}{V}X_r - \frac{F}{V}(1 + r)X + \mu X \qquad (24.36)$$

F = flow rate, liter/hr
V = volume, liter
X = cell conc., g/liter
S = substrate conc., g/liter
r = recycle ratio
C = cell conc. ratio

Fig. 24.9. Single-stage continuous fermentor with recycle.

from which:

$$\mu = D\left(1 + r - r\frac{X_r}{X}\right) = D(1 + r - rC) \tag{24.37}$$

Because $(1 + r - rC) < 1$, it is possible to operate the system at dilution rates greater than the maximum growth rate. Utilizing the Monod equation, the following expressions for the growth limiting substrate, S, system effluent biomass concentration, X_e, and biomass concentration in the exit stream from the bioreactor, X, are obtained:

$$S = \frac{K_s D(1 + r - rC)}{\mu_{max} - D(1 + r - rC)} \tag{24.38}$$

$$X_e = Y(S_0 - S) \tag{24.39}$$

$$X = \frac{Y(S_0 - S)}{(1 + r - rC)} \tag{24.40}$$

These relationships are illustrated in Fig. 24.10 for a particular set of conditions.

Another continuous fermentor system encountered in the microbial world is the multistage system with step feeding. This configuration is depicted in Fig. 24.11.

In this system, fresh feed is step-fed, i.e., fed to all reactor stages. This configuration is used in waste treatment systems to provide greater stability and to minimize the effects of substrate "shock loading" or inhibition through distribution of the concentrated feed along the system. A two-stage step feeding system also is used as a research tool to look at the inhibitory effects of high substrate loading which cannot be achieved in the single stage without washout. The material balance relationships for step feeding systems are straightforward. What is usually troublesome is the selection of the proper kinetic expression for high substrate conditions.

Fig. 24.10. Relationships for a single-stage continuous fermentor with recycle.

Fig. 24.11. Multistage continuous fermentor with step feeding.

Novel Bioreactors

The 1980s witnessed some innovative bioreactor designs and applications, largely stimulated by the need to commercialize therapeutic proteins whose production in large quantity was made possible by recombinant DNA and hybridoma technologies. Particularly worth mentioning are hollow fiber perfussion reactors used for the production of secreted products such as monoclonal antibodies, microcarrier culture for growth of anchorage dependent cells, and *in situ* rotating filters (shelves of screens) used for retention of cells and microcarriers and exchange of medium and sometime gases.

Oxygen Transfer in Fermentation Systems

In aerobic fermentations, oxygen is a basic substrate that must be supplied for growth. As in enzymatic reactions, the relationship between oxygen concentration and growth is of a Michaelis-Menton type. The specific rate at which cells respire (Q_{O_2}) increases rapidly with an increase in the dissolved oxygen concentrations (C) up to C_{crit}, which typically ranges from 0.5 to 2.0 ppm for well-dispersed

bacteria, yeast, and fungi growing at 20 to 30°C. Beyond C_{crit}, the specific oxygen uptake increases only slightly with increasing oxygen concentrations.

If Q_{O_2} is plotted against the specific growth rate (μ) of the microbe, a linear correlation is obtained. An intersection of this straight line with the ordinate is designed as $(Q_{O_2})_m$, which represents the oxygen required for cellular maintenance.

In mathematical terms, this relation is given by

$$Q_{O_2}X = (Q_{O_2})_m X + (1/Y_{X/O_2})(dX/dt) \quad (24.41)$$

where Y_{X/O_2} is the yield based on cell mass, and t is the time.

This type of correlation can be applied to almost any substrate involved in cellular energy metabolism. It is justified from an energetic point of view and is supported by experimental steady-state data. Care must be used in applying this equation to transient conditions. It can be used only for situations at or near the steady-state equilibrium conditions.

When a product other than cell biomass is involved, the situation can be considerably more complex. Oxygen can be utilized for maintenance and growth as well as product formation. The simplest possible expression to represent the rate of oxygen utilization in this case is

$$Q_{O_2}X = (Q_{O_2})_m X + (1/Y_{X/O_2})(dX/dt) + (1/Y_{P/O_2})(dP/dt) \quad (24.42)$$

where P is the product concentration.

Of course, the oxygen transfer rate under steady-state conditions must be equal to the oxygen uptake rate, i.e.:

$$Q_{O_2}X = k_L a(C^* - C)_{mean} \quad (24.43)$$

where C^* = concentration of oxygen in the liquid that would be in equilibrium with the gas-bubble concentration, and $k_L a$ = the volumetric oxygen transfer rate.

Utilizing this relationship, the volumetric oxygen transfer coefficient can be estimated, i.e.:

$$k_L a = (Q_{O_2}X)/(C^* - C)_{mean} \quad (24.44)$$

Fig. 24.12. Determination of oxygen consumption and mass transfer rate by dynamic techniques.

In practice, the dissolved oxygen concentration (C) is monitored by a membrane-covered probe. When the fermentation is equipped with a fast-responding probe (i.e., the response time is less than 6 sec), then the fermentor can be used as a respirameter by making dynamic measurements of the oxygen concentration under aeration and nonaeration conditions. This is illustrated in Fig. 24.12.

The following equations represent these conditions. Air off, Phase I:

$$dC/dt = -Q_{O_2}X \quad (24.45a)$$

Air on, Phase II:

$$dC/dt = K_L a(C^* - C)_{mean} - Q_{O_2}X \quad (24.45b)$$

Once $Q_{O_2}X$ is obtained from dC/dt in the "air-off" phase, $k_L a$ can be calculated from the dissolved oxygen profile of the "air-on" phase.

Scale-up

Many fermentation companies have the problem of scaling-up new fermentations as well as the translation of process-improvement data for well-established fermentations from laboratory operations to existing plant equipment. In general, fermentations are scaled up on the basis of achieving similar oxygen transfer capabilities in the plant equipment that proved to be optimal on the bench scale.

Fermentation biomass productivities usually range from 2 to 5 g/liter/hr. This represents

an oxygen demand in the range of 1.5 to 4 g O_2/liter/hr. In a 500 m^3 fermentor this means achievement of a volumetric oxygen transfer coefficient in the range of 250 to 400 hr^{-1}. Such O_2 transfer capabilities can be achieved with aeration rates of the order of 0.5 VVM and mechanical agitation power inputs of 3.3 to 4.4 hp/m^3 (1.2 to 1.6 hp/100 gal).

On scale-up, however, it usually is heat removal that causes design problems. With the above mechanical agitation power inputs, between 40 and 50 Btu/gal/hr of turbulent heat is generated. This, coupled with a peak metabolic heat release from the growing biomass of 120 to 180 Btu/gal/hr, means that the fermentor must be capable of removing up to 160 to 230 Btu/gal/hr. If the fermentation is a penicillin fermentation operated at 25°C, and if the best cooling water temperature that can be achieved is around 18°C, and if the overall heat transfer coefficient is in the range of 100 Btu/hr/°F/ft^2, there is no way that the heat will be removed in large fermentors without external heat exchange or extensive cooling coils in the fermentor. In highly viscous fermentations, internal cooling coils usually are not desirable because of interference with mixing patterns.

As a consequence, numerous schemes exist for heat removal in large fermentors including half coil baffles and draft tubes. No common design seems to have emerged. Manufacturers each have their own design scheme that they promote.

Power input to the liquid phase in a sparged system (P_g) can be calculated from:

$$P_g/V = \frac{Q_g P_g}{V}\frac{\eta v_o^2}{2} + \frac{RT}{M}\left(ln\frac{\pi_o}{\pi}\right) \quad (24.46)$$

The first term represents the energy dissipated at the sparger holes. The second term represents the energy involved to move the gas through the static liquid head. For well-designed spargers the first term usually can be neglected.

Unaerated mechanical power input (P_{mo}) can be estimated from:

$$P_{mo}/V = \frac{f(\rho N^3 D^5)}{V} \quad (24.47)$$

The power factor in this correlation can be obtained from the paper of Rushton et al.[46]

For estimation of mechanical power input to a sparged agitated system (P_m), the correlations of Michel and Miller are widely used:

$$P_m/V = \frac{0.706}{V}\left[\frac{P_{mo}^2 N D^3}{Q_g^{0.56}}\right]^{0.45} \quad (24.48)$$

Although the separate effects of sparged and of mechanically agitated power input can be estimated, their effect is not simply additive. In general, a correlation of the following type is used to estimate their combined effect:

$$\frac{P_e}{V} = \frac{P_m}{V} + \frac{CP_g}{V} \quad (24.49)$$

The C correlation term usually varies from 0.05 to 0.4 in value. The data of Miller[47] can be used to estimate the value of the constant.

There are various correlations between k_La and power inputs. Some design engineers prefer to scale-up on the following basis:

$$k_La = k(P_m/V)^\alpha (v_s)^\beta \quad (24.50)$$

where $\alpha = 0.95$ and $\beta = 0.67$ for pilot-plant equipment and $\alpha = 0.4$ and $\beta = 0.5$ for large-scale plant equipment. When such a basis is used for scale-up, the designer attempts also to maintain constant tip speed on scale-up. Generally, these run from 600 to 1200 ft/min in most fermentors, with 1000 ft/min a typical value.

Once a plant is built, the conditions of agitation, aeration, mass (oxygen) transfer, and heat transfer are more or less set. Therefore, it has been suggested that the problem of translating process improvements is one not of scale-up but rather of scale-down. Those environmental conditions achievable in plant-scale equipment should be scaled down to the pilot-plant and screen-size equipment (shake-flask) to ensure that the studies are carried out under conditions that can be duplicated.

Scale-up of mammalian cells creates different dimensions of challenges for engineers. First of all, most mammalian cells with their lack of protecting cell wall are very sensitive to shears that are generated with high-speed

agitation. Some of the cells are so sensitive to shear that they are damaged upon the collapse of gas bubbles. Sparging, a common aeration mechanism used in the fermentation industry, thus becomes unsuitable for "aerating" these cells. So, how should one go about mixing and gassing mammalian cell cultures? Low-shear marine propellers or pitched blades generally are used to provide gentle mixing. A hemispherical, instead of the conventional, dished bottom, sometimes is incorporated into the bioreactor design. Gas-permeable silicone tubing is used for "bubbleless" aeration or oxygenation. And the rotating filters (shieves), mentioned earlier, used for cell retention and medium exchange, can be conveniently used as a gassing cage. An external loop including an oxygenation chamber is just another means of introducing oxygen to the system without directly sparging on the cells.

Air Sterilization

Submerged aerobic fermentation processes require a continuous supply of large quantities of air. Sterilization of this air is mandatory in many fermentations. For pure culture operation, incomplete destruction or inadequate removal of the microorganisms carried in the air may preclude successful operation. Many ways have been suggested for sterilizing air. Only filtration through beds of fibrous and membrane-type materials have found widespread usage on an industrial scale. Examples would be filtration through beds of fibrous materials such as glass wool or plate-type filters made of polyvinyl alcohol several centimeters thick. In recent years sufficient research has been carried out to permit the design of these fibrous filters on a rational basis. Major requirements that every air sterilization system must satisfy are as follows:

1. The system should be simple in design.
2. It should not be inordinately costly to operate.
3. It should remove or destroy airborne contamination to the extent necessary for satisfactory fermentation performance.
4. It should be stable to repeated steam or chemical vapor sterilization.
5. It should condition the air.
6. Its ability to maintain a sterile air supply should not be jeopardized by power failure or compressor surges.

This last requirement frequently is overlooked. Its consideration is paramount in the design of filters compounded from fibrous materials. For a particular filter there is an intermediate air velocity at which filtration effciency is a minimum. If the filter design is based upon a performance observed at an operating velocity other than that at which minimum efficiency occurs, surges or brief power failures could create periods of operation at lower than designed-for efficiencies. Hence the filter design should be based on satisfactory operation at this point of minimal efficiency.

Minimal efficiency at an intermediate air velocity occurs because different forces act to collect airborne particles at different velocities. At low velocities, gravitational, diffusional, and electrostatic forces act on the particle. Their effect is inversely proportional to air velocity. At high velocities, inertia forces come into play, and they are directly proportional to air velocity. The nature of inertial effects is such that below a certain air velocity, collection due to inertial forces is zero.

This velocity can be estimated by the following relation:

$$V_{\text{min eff}} = \frac{1.125\mu d_f}{C\rho_p d_p^2} \qquad (24.51)$$

where

μ = air viscosity
d_f = fiber diameter
C = Cunningham correction factor
ρ_p = particle density
d_p = particle diameter

For the collection of unit-density, one-micron-sized bacterial particles from air streams at room temperature and pressure. This velocity is equal to

$$V_{\text{min eff}} = 0.066 d_f \qquad (24.52)$$

where velocity V is in feet per second and the fiber diameter, d_f, is in micrometers.

Regardless of air velocity, some collection always occurs because airborne particles possess a finite size and will be intercepted by some fiber blocking an air stream along which a particle moves.

Collection must always be greater than that due to interception because it represents the minimum collection physically possible. In the absence of experimental data, a reasonable estimate of the minimal collection efficiency of a particular filter is that estimated solely from interception effects calculated at a velocity just below which inertia effects are zero.

The theoretical expression for the collection efficiency of a single isolated fiber, due solely to interception effects, can be estimated from

$$n_o = 0.5[1/(2 - lnN_{Re})][2(1 + R)ln(1 + R) - (1 + R) + 1/(1 + R)] \quad (24.53)$$

where

n_o = collection efficiency of a single isolated fiber
$R = d_p/d_f$
N_{Re} = Reynolds number

From experimentation with aerosols it has been found that the effectiveness of a filter can be expressed by

$$\ln\frac{N_1}{N_2} = \frac{1.27n_o(1 + 4.5\alpha)\alpha L}{(1 - \alpha)d_f} \quad (24.54)$$

where:

N_1 = total number of particles entering the filter
N_2 = total number of particles penetrating the filter
L = filter thickness
α = volume fraction of fibers in the filter

This equation is a reasonable basis for air sterilizing filter design. The single fiber efficiency, n_o, to use in this equation when experimental data are lacking is that based on collection due solely to interception effects estimated at an air velocity where minimal filtration efficiency is expected.

Although depth filters are still in use, because of the interest in maintaining absolute asepsis particularly for long-term culture and for ease of operation, membrane-type absolute filter cartridges are becoming more popular. In this case, a prefilter for removal of particulates and liquid droplets usually is placed upstream of the absolute filter. Parallel installation of the filters prevents a total shutdown of the fermentation in the event of filter clogging.

Medium Sterilization

Microorganisms can be removed from fluids by mechanical methods, for example, by filtration, centrifugation, flotation, or electrostatic attraction. They also may be destroyed by heat, chemical agents, or electromagnetic irradiation. Although cells may be disrupted or killed by mechanical abrasion on a small scale, this method is not satisfactory industrially. Similarly, X rays, β rays, ultraviolet light, and sonic irradiations, although useful in the laboratory, are not applicable to the sterilization of large volumes of fluids.

Although antibacterial agents have an important place in the fermentation industry, particularly for the production of a pure water supply, they have little application for the sterilization of fermentation media. Despite the fact that heat sterilization of media is the most common method, little attention has been paid until recently to the engineering aspects of heat sterilization.

Interest in continuous methods of sterilizing media is increasing, but for the successful operation of a continuous sterilizer, foaming of the media must be carefully controlled and the viscosity of the media must be relatively low. Figure 24.13 illustrates two types of continuous media sterilizers that have been utilized in the fermentation industry. The advantages of continuous sterilization of media are as follows:

1. Increase of productivity because the short period of exposure to heat minimizes damage to media constituents.
2. Better control of quality.

3. Leveling of the demand for process steam.
4. Suitability for automatic control.

At present, most media in the fermentation industry are sterilized by batch methods. Overexposure of the medium to heat is inherent in batch sterilization processes. But continuous sterilization, when properly operated, can minimize damage to the medium.

Design and operation of equipment for sterilizing media are based on the concept of thermal death of microorganisms. Consequently, a understanding of the kinetics of the death of microorganisms is important to the rational design of sterilizers.

The destruction of microorganisms by heat implies loss of viability, not destruction in the physical sense. The destruction of organisms by heat at a specific temperature follows a monomolecular rate of reaction:

$$\frac{dN}{dt} = -KN = -(Ae^{-E/RT})N \quad (24.55)$$

Fig. 24.13. Two types of continuous sterilizers.

where:

k = reaction rate constant, time^{-1}
N = number of viable organisms/unit volume
t = time
T = absolute temperature
E = energy of activation for death
R = gas law contact
A = constant

This equation can be integrated to yield the design equation:

$$\ln \frac{N_0}{N_f} = A \int_0^{t_s} e^{-E/RT} dt \quad (24.56)$$

where:

N_0 = number of contaminating organisms in the total fermentation medium to be sterilized
N_f = level of contamination that must be achieved to produce the desired degree of apparent sterility
t_s = sterilization time

In estimating the medium sterilization time, one must define the contamination, the desired degree of apparent sterility, and the time–temperature profile of the medium, i.e., $T = f(t)$. For typical bacterial spore contaminants, the constants used in most designs have the following values:

$$E = 68{,}700 \text{ cal/g-mol}$$

$$R = 1.987 \text{ cal/g-mol, °K}$$

$$A = 4e^{+87.82}, \text{min}^{-1}$$

The medium used for culturing mammalian cells must be filter-sterilized. For mammalian cell culture where viral attacks are more of a concern than for microbial fermentations, filters of 0.45 micron and 0.1 micron, placed in series, commonly are used.

Instrumentation and Control

In scaling up any successful fermentation, the molecular biologist would suggest the necessity for controlling the environment and hence regulating the fermentation. However, two problems arise. First, knowledge may not exist

of the regulatory mechanisms of metabolic pathways that produce the desired product. In fact, the metabolic pathways may not be fully known or understood. Second, even if the pathways and the regulatory mechanisms are known, the necessary instrumentation to detect regulatory metabolites may not exist. The design engineer thus is presented with a dilemma in developing a new fermentation or improving an existing process.

Should the engineer focus on researching the mechanisms of product regulation and control, developing needed analysis and sensing instrumentation to provide this control, or should he or she use trial-and-error development procedures for strain and medium selection to evolve an apparent optimal environment for plant-scale production? The answer is that the engineer must do both.

In the past, metabolic controls simply were bred into or out of fermentation organisms through mutation and strain selection. However, tremendous strides have been, and are being, made in sensor development. The design engineer will soon be able to reply more on environmental control than before in order to gain economical fermentation results. Figure 24.14 diagrams show a highly instrumented fermentor, which is designed to secure basic information on environmental control,

Fig. 24.14. General layout of a highly instrumented, computer-coupled fermentor.

can be coupled to a computer for data analysis and control.

To achieve meaningful fermentation control, the following three steps are necessary:

1. Carry out fermentation research on fully monitored environmental systems.
2. Correlate the environmental observations with existing knowledge of cellular control mechanisms.
3. Reproduce the desired environmental control conditions through continuous computer monitoring, analysis, and feedback control of the fermentation environment.

Until recently, fermentation control was limited to that of temperature, pH, and aeration. With the development of numerous sensors and inexpensive personal and minicomputers, the engineer can think in terms of sophisticated control systems for fermentation processes.

Indeed, the engineer can even think of utilizing computers to perform on-line dynamic optimization.

New instruments still are needed. The most important sensor needed is a reliable biomass monitoring device that can be sterilized.

An IR fiber-optic cell density probe has enjoyed some success. It has been used to monitor cell growth directly (without dilution) in high-cell-density bacterial fermentations. Still, there is a need for a glucose, a nitrogen substrate, and a phosphate sensor that can withstand repeated system sterilization. However, the ability to do on-line sample filtration through the use of hollow fibers or rotating filters has made possible the continuous, on-line measurement of glucose, lactate, and so on.

Indirect measurement via computers appears to be a viable alternative to measurement involving sampling. Certain sensor information can be combined to give additional information such as oxygen-uptake rate, carbon-dioxide evolution rate, and respiratory quotient. These measurements can be thought of as "gateway" measurements because they make possible the calculation of additional information (Table 24.3).[1]

TABLE 24.3 Gateway Measurements*

Measurement	*Result*
pH	Acid product formation rate
Air flow rate	O_2 uptake rate
In and out O_2 concentration	
Air flow rate	CO_2 evolution rate
Out CO_2 concentration	
CO_2 evolution rate	Respiratory quotient
O_2 uptake rate	
Power input	O_2 transfer rate
Air flow rate	

*S. Aiba, A. E. Humphrey and N. Millis, *Biochemical Engineering*, 2nd ed., p. 332, University of Tokyo Press, Tokyo (1973).

Further, the indirect measurement of a given component can be achieved by material-balancing that component around the fermentor. If a model for utilization of that component for biomass or product formation is known, then either the biomass or the product level can be estimated by computer summation or integration of the data using the model.

Besides these uses, the computer has application in fermentation processes for continuous nonprejudicial monitoring and (most important) continuous feedback control and dynamic optimization of the process.

Recovery of Fermentation Products

From the amount of space devoted to fermentor design and scale-up, one might gather that the recovery processes of fermentation products are rather straightforward and relatively simple. Nothing could be further from the truth. In one case of an

Fig. 24.15. Basic flowsheet for the recovery of antibiotics. (Ind. Eng. Chem., *49, 1494, 1957. Copyright American Chemical Society.*)

Fig. 24.16. Basic flowsheet for an enzyme production plant. (Process Biochem., **5**, 22, *1970*.)

antibiotic production plant, the investment for the recovery facilities is claimed to be about four times greater than that for the fermentor vessels and their auxiliary equipment. As much as 60 percent of the fixed costs of fermentation plants are attributable to the recovery portion in organic acid and amino acid fermentations.

Figure 24.15 shows a typical recovery process for antibiotics, and Fig. 24.16 presents a flowsheet for an enzyme plant. It is apparent from these diagrams that most recovery processes involve combinations of the following procedures:

1. Mechanical separations of cells from fermentation broth.
2. Disruption of cells.
3. Extraction.
4. Preliminary fractionation procedures.
5. High-resolution steps.
6. Concentration.
7. Drying.

Chromatography technology has advanced rapidly in the last decade. Various separation principles, gel filtration, size exclusion, ion exchange, hydrophobic interaction, and affinity binding have been applied in the chromatographical mode. Many different pressure-sustaining support matrices have been developed for high-throughput applications. All of these techniques have greatly expanded the tools now available for high-resolution separation.

Immobilized Enzymes and Cells

Most of the soluble hydrolytic enzymes in common use in industry are formed extracellularly; the wide range of intracellular microbial enzymes are virtually unexplored commercially. In order to exploit these enzymes, it is necessary to develop economic methods of purification. Reports of continuous methods for harvesting, breaking cells, and fractionating their protein represent major advances toward this end. Recently, enzymes as well as whole cells have been imobilized by adsorption, encapsulation, or inclusion in gels. They also may be covalently bound with a bifunctional linking agent to an insoluble polymer, covalently cross-linked with themselves, or bound directly in an ezyme–polymer complex. The preparation can be packed into a column or complexed onto porous sheets, and the substrate then can react with the enzyme in a batch or a continuous process. Enzymes can also be contained within an ultrafiltration membrane. For instance, starch can be fed continuously into a membrane structure containing α-amylase and the products of hydrolysis collected outside the membrane. Whole cells or crude extracts also can be placed in capsules that allow substrate and products to diffuse to the enzymes.

Some properties of the enzyme, such as the pH optimum, Michaelis constant, or stability, may change when it is immobilized, but the economies achieved by retaining the enzyme for reuse and obtaining products free of contaminating enzyme greatly outweigh any loss of stability or reduction in the rate of reaction. In many cases the life of the immobilized enzyme is greater than that of the soluble enzyme. If substrates or products are sensitive to pH, it may be possible to select a polymer that changes the optimum pH of the immobilized enzyme to a range more favorable to the stability of the reactants.

ORGANIC ACIDS

There are many organic acids that can be produced by microbial or biochemical means. The major ones are listed in Table 24.4. However, only acetic acid (as vinegar), citric acid, itaconic acid, gluconic acid, and 2-ketogluconic acid are produced industrially by ferentation now. Other organic acids, such as fumaric, gallic, lactic, malic, and tartaric acids, once produced by fermentation or enzyme processes, now are produced commercially predomiantly by the more economic means of chemical synthesis.

Acetic Acid and Vinegar

Acetic acid as a chemical feedstock is manufactured by chemical synthesis. Acetic acid in

TABLE 24.4 Major Organic Acids That Can Be Produced by Fermentation

Organic Acids	Producing Microorganisms	*Substrates*	*Yields %*
Acetic	*Acetobacter acetic*	Ethanol	95
	Clostridium thermoaceticum	Glucose	90
Araboascorbic	*Penicillium notatum*	Glucose	45
Citric	*Aspergillus niger*	Sucrose	85
	Candida lipolytica	*n*-Paraffin	140
Fumaric	*Rhhizopus delemar*	Glucose	58
Gluconic	*Aspergillus niger*	Glucose	95
Isocitric	*Candida brumptii*	Glucose	28
Itaconic	*Aspergillus terreus*	Glucose	60
2-Ketogluconic	*Pseudomonas fluorescens*	Glucose	90
5-Ketogluconic	*Gluconobacter suboxydans*	Glucose	90
α-Ketoglutaric	*Candida hydrocarbofumarica*	*n*-Paraffin	84
Kojic	*Aspergillus oxyzae*	Glucose	50
Lactic	*Lactobacillus delbriickii*	Glucose	90
Malic	*Lactobacillus brevis*	Glucose	100
Propionic	*Propionibacter shermanii*	Glucose	60
Pyruvic	*Pseudomonas aeruginosa*	Glucose	50
Salicylic	*Pseudomonas aeruginosa*	Naphthalene	94
Succinic	*Bacterium succinicum*	Malic acid	57
Tartaric	*Gluconobacter suboxydans*	Glucose	27
Xylonic	*Enterobacter cloacae*	Xylose	90

the form of vinegar (at least 4% acetic acid by law) is produced largely via the oxidation of ethanol by bacteria of the *Acetobacter* genus.

Synthetic acetic acid is presently produced by several different routes:

1. Wood distillation (being phased out).
2. Oxidation of acetaldehyde (Wacker process).
3. Liquid-phase oxidation of *n*-butane.
4. Carbonylation of methanol (Monsanto process).

The last method is the most economical among the four, and is employed in most new acetic acid manufacturing facilities throughout the world. Monsanto was awarded the Kirkpatrick Chemical Engineering Merit Award in 1976 for the development of this process.

The U.S. annual production of acetic acid was about 3.4 billion lb in 1987, and the selling price around $0.30/lb. The world production totaled 10 billion lb. The major U.S. producers are Celanese, Monsanto, USI, Union Carbide, Eastern, and Borden.

High-tonage production of acetic acid for industrial purposes was earlier done through fermentation. In 1915, USI of Baltimore, Maryland, constructed and operated two of the largest acetic acid fermentation plants ever built, with a combined tank capacity of about 20 million gallons. The acetic acid was converted with lime to calcium acetate, and the latter pyrolyzed to acetone. The plant operated for a number of years until competition from acetone produced by direct fermentation made the process uneconomical. Today, fermentive acetic acid is limited to vinegar production only.

Vinegar is one of the oldest known fermentation products, predated only by wine and possibly by certain foods from milk. First derived from the spoilage of wine, vinegar has been used as a condiment, food preservative, medicinal agent, primitive antibiotic, and even today a household cleansing agent. The present production of vinegar is used almost entirely in foods.

Vinegar may be defined as the product of a double fermentation: an alcoholic fermentation of a sugary mash by a suitable yeast (usually a selected strain of *Saccharomyces cerevisiae* or *ellipsoidens*) and a second fermentation to oxidize the alcohol to acetic acid by a suitable culture of *Acetobacter* organisms.

The theoretical maximum of acetic acid yield on glucose is 67 percent (two moles of acetic acid produced from every mole of glucose consumed) by this route. A homofermentative culture, *Clostridium thermoaceticum*, is known to be capable of fixing CO_2 and yielding three moles of acetic acid from one mole of glucose under anaerobic conditions. The technology for this process has not been commercialized, however.

Vinegar can be made from a variety of fermentable substances, the essential requirements being that they are satisfactory and economical sources of ethanol. The commonly used substances include fruits and their juices, cereals, sugar syrups, and synthetic ethanol. There are four major types of vinegar: white distilled, cider, wine, and malt vinegar. Almost the entire U.S. production of white distilled vinegar is derived from synthetic ethanol. Cider vinegar is made by the alcoholic and subsequent acetous fermentation of the juice of apples or its concentrate; wine vinegar is made of grape wine; and malt vinegar made of a mash containing malt, corn, and/or barley. The annual U.S. production of all types of vinegar amounts to about 150 million gallons, approximately 77 percent being white distilled, 17 percent cider, 4 percent wine, and 1 percent malt vinegar. The major producers include Heinz U.S.A., Hunt Foods, National Vinegar, Speas, and Standard Brands.

Several vinegar manufacturing processes are commercially used, including the following:

1. Circulating generators (trickling generators).
2. Frings Acetators (submerged culture generators).
3. Yeomans Cavitators.
4. Tower fermentors (column fermentors).

The circulating, trickling generator is most widely used. It is a large tank constructed in a variety of dimensions, generally of wood, including redwood and fir but preferably cypress. The vertical timbers are held in place with steel hoops. A false bottom supports curled beachwood shaving above the lower one-fifth of the tank, which serves as a collection reservoir. Air is supplied by a simple fan-type blower, and is distributed to the generator by a number of equally spaced inlets just beneath the false bottom. An air flow rate of about 0.015 vvm (volume of air per volume of packing per minute) is adequate. A pump circulates the ethanol–water–acetic acid mixture from the collection reservoir up through a cooler to a distributing sparger arm in the top of the tank. The liquid trickles down through the packing and returns to the bottom reservoir. Cooling water to the cooler is regulated to maintain the temperature of the generator around 29°C at the top and below 35°C at the bottom. A portion of the finished vinegar is periodically withdrawn from the reservoir, and replaced with the ethanol-containing charge. The ethanol concentration in the generator should not exceed 5 percent or fall below 0.2 percent. If ethanol is depleted in the generator, the *Acetobacter* will die, and the generator becomes inactive.

The Frings Acetator (produced by the Heinrich Frings Company of Bonn, Germany) consists of a stainless steel tank with internal cooling coils, a high-speed, bottom-entering agitator, and a centrifugal foam breaker. The unique feature of this Acetator is its highly efficient method of supplying air. This is accomplished by means of a high-velocity self-aspirating rotor that pulls air in from the room to the bottom of the tank. The equipment is operated batch-wise. When the ethanol content falls to 0.2 percent by volume, about 35 to 40 percent of the finished product is removed. Fresh feed is pumped in to restore the original level, and the cycle starts again. Cycle time for 12 percent vinegar is about 35 hours. The rate of production can be as much as ten times as great as that obtained in the trickling generator of equivalent size. The

yield on ethanol is higher. Values of 94 percent and 85 percent have been reported, respectively, for the Acetator and the trickling generator. But much more extensive refining equipments are necessary for filtering vinegar produced by the submerged process because the mash contains the bacteria that produced it.

The Yeomans Cavitator (no longer being manufactured) is a submerged culture system somewhat similar to the Acetator but differing in the way in which air is supplied. This generator has a top-driven rotor-agitator. It withdraws liquid and air from a centrally located draft tube. The system can be easily installed in existing wooden vats and adapted to continuous production. Technical difficulties have forced its abandonment. Some units are still in use to produce vinegar is the continuous mode. A 98 percent efficiency of ethanol to acetic acid has been achieved in commercial operations.

The tower fermentor is a relatively new aeration system applied to vinegar production. The fermentor is constructed of polypropylene reinforced with fiber glass. Aeration is accomplished through a plastic perforated plate covering the cross section of the tower and holding up the liquid. The cost of the tower fermentor is said to be approximately half that of a Frings Acetator of equivalent productive capacity. It has been reported that the tower fermentor is satisfactory for producing all types of vinegar.

Vinegar clarification is accomplished by filtration, usually with the use of filter aids and/or fining agents such as diatomaceous earth or bentonite, respectively. After clarification, vinegar is bottled, sealed tightly, pasturized at 60 to 65°C for 30 minutes, and then cooled to 22°C. Vinegar can be concentrated by a freezing process. Vinegar of 200-grain strength is readily obtainable from 120-grain raw vinegar. There are a number of advantages of 200-grain vinegar. The acid strength of brine solutions that would have had to be discarded because of dilution by pickle juice can easily be increased by pickle processors. Transportation costs are substantially reduced, as well.

Nakano Vinegar is the largest manufacturer of food vinegar in Japan. The company is famous for its brand name Mitukan. Nakano, in collaboration with Professor T. Beppu, University of Tokyo, has cloned alcohol dehydrogenase and aldehyde dehydrogenase, and established phenotypic transformation of *Acetobacter aceti*. The acetic acid bacteria with the aldehyde dehydrogenase gene produced acetic acid more rapidly than those with the alcohol dehydrogenase, and were more resistant to high concentrations of acetic acid.

Citric Acid

Citric acid, whose structure is shown below, is the most important organic acid produced by means of fermentation.

In the food and beverage industries, citric acid is used in soft drink mixes, in carbonated and still beverages, candies, wines, desserts, jellies, and jams, as an antioxidant in frozen fruits and vegetables, and as an emulsifier in cheese. These uses represent 64 percent of the use of citric acid. As the most versatile food acidulant, citric acid accounts for about 75 percent of the total food acidulant market.

In terms of pharmaceutical and cosmetic applications, citric acid provides effervescence in oral dosage forms and in products for external use by combining the citric acid with a biocarbonate/carbonate source to form carbon dioxide. Citric acid and its salts are also used in blood anticoagulants to chelate calcium, block blood clotting, and buffer the blood. Citric acid is contained in various cosmetic products such as hair shampoos, rinses, lotions, creams, and toothpastes. These areas of application account for 20 percent of citric acid use.

More recently, citric acid is being used for metal cleaning, in detergents substituting phosphate, as plasticizers in its ester forms, for secondary oil recovery, and as a buffer/absorber in stack gas desulfurization. The use

of sodium citrate in heavy-duty liquid laundry detergent formulations is responsible for citric acid's rapid increase in use.

Citric acid was first isolated from lemon juice and crystallized as a solid in 1784 by Scheele. In 1893, Wehmet first described citric acid as a product of mold fermentation. In 1919, fermentation processes based on sucrose were developed commercially. Many organisms have been shown to produce citric acid from carbohydrates as well as from *n*-paraffins; however, *Asperigillus niger* has always given the best results in industrial production of citric acid.

The generally accepted pathway from sugar to citric acid is assumed to follow the glycolysis pathway to pyruvate. Pyruvate then enters the TCA cycle. It is decarboxylated to acetate as acetyl-CoA, or adds on CO_2 to form oxalacetate. Acetyl-CoA and oxalacetate then react to form citrate. It is generally considered that aconitase, the enzyme responsible for the conversion of citrate to isocitrate in the TCA cycle, is inhibited by depriving it of the necessary metal (iron) coenzymes. Therefore, the TCA cycle does not operate, and citric acid accumulates. Figure 24.17 summarizes the reactions leading to citric acid from glucose. It is worthwhile to note that one mole of glucose yields one mole of citric acid with no consumption of oxygen. The overall reaction is actually energy-yielding. It yields one mole of ATP and two moles of $NADH_2$ per mole of citric acid produced. Both ATP and $NADH_2$ are high energy-containing molecules. This aspect makes the

Fig. 24.17. Pathway leading to citric acid from glucose.

process a good candidate for the process of cells immobilization.

Microbiological production of citric acid can be implemented by three techniques:

1. Solid state fermentation.
2. Liquid surface fermentation.
3. Submerged culture fermentation.

In solid state or Koji fermentation, *Aspergillus niger* grows on moist wheat bran (70–80 percent water) and produces citric acid in 5 to 8 days. This process is practiced only in Japan and accounts for about one-fifth of Japanese citric acid production.

In liquid surface or shallow tray fermentation, beet molasses (containing 48–52% sugar) or cane molases of blackstrap (containing 52–57% sugar) or high test (containing 70–80% sugar) is introduced into a mixer. Dilute sulfuric acid is added to adjust the pH to about 6.0. Phosphorus, potassium, and nitrogen in the form of acids or salts are added as nutrients for proper mold growth and optimal citric production. The mix then is sterilized and finally diluted with water to a 15 to 20 percent sugar concentration. The medium flows by gravity into shallow aluminum pans or trays arranged in tiers in sterile chambers. Most chambers have provisions for regulation and control of temperature, relative humidity, and air circulation. One plant has 80 trays per chamber. Each tray holds about 400 liters of solution at a depth of 76 mm. When the medium has cooled to about 30°C, it is inoculated with spores of *Aspergillus niger*. The tray fermentation requires 8 to 12 days. The pH drops to about 2 at the end of the fermentation, and the acid content varies from 10 to 20 percent. Some oxalic and gluconic acids also are formed. The temperature is maintained at 28 to 32°C during the fermentation. Sterile air is circulated through the chambers, and the relative humidity is controlled between 40 and 60 percent.

The submerged culture or deep fermentation process has been adopted by most newly constructed plants. The fermentation medium consists of sucrose (around 200 g/liter) and mineral salts to provide a balanced supply of iron, zinc, copper, magnesium, manganese, and phosphate. The provision of a suitable culture medium is the most critical factor in obtaining a high yield of citric acid. The fermentation is carried out at 25 to 27°C. Continuous aeration is provided by bubbling air at a rate of 0.5 to 1.5 volumes of air/volume of solution/minute. Mechanical stirring is not necessary. It is generally accepted that the formation of pellets between 1 and 2 mm in diameter in the fermentation mash is most desirable. Pelleting reduces both viscosity, increases oxygen transfer, and simplifies mycelium separation in the recovery scheme. The submerged fermentation has a time cycle of 6 to 9 days.

The yield of citric acid on sugar varies from process to process and from manufacturer to manufacturer. The theoretical maximum is 112 percent on sucrose. The liquid surface fermentation has a yield of 90 to 95 percent, and the submerged culture fermentation 75 to 85 percent. Improvements have been made by reducing the formation of by-products, mainly oxalic acid, and yields of the submerged culture process are reaching those of the surface culture process.

The fermentation broth from the solid state, surface culture, or submerged culture process is treated similarly for recovering and refining citric acid. Two recovery methods are being used: (1) precipitation and filtration and (2) extraction.

A process flowsheet including the fermentation section and the refining section using the first method is shown in Fig. 24.18. The mycelium is filtered out of the fermentation liquor first. The mycelium may be used as fertilizer after proper weathering and processing. The clarified liquor flows to precipitating tanks fitted with stirrers, where it is heated to a temperature of 80 to 90°C. The oxalic acid present is separated by preferential precipitation through the addition of a small amount of hydrated lime. The resulting calcium oxalate is worked up separately in a manner similar to the following process described for citric acid recovery. Approximately one part of hydrated lime for every two parts of liquor is added slowly over a one-hour period while the temperature is raised to about 95°C. The

Fig. 24.18. Citric acid process flowsheet.

precipitated calcium citrate is filtered on a vacuum filter, and the filtrate free of citrate is run to waste. The calcium citrate cake is run to acidulation tanks, where it is acidified with dilute sulfuric acid. Then it is filtered, and the citric acid mother liquor is decolorized by a charcoal treatment. The purified liquor next is concentrated in a vacuum evaporator, and then it is run into a crystallizer where, on cooling, citric acid crystallizes, generally in the form of the monohydrate. The resulting acid is of USP grade.

The extraction method treats the filtered fermentation liquor with a highly selective solvent, tri-*n*-butyl phosphate, and then recovers free citric acid by counterextraction with water. The aqueous solution, which is further concentrated and crystallized, yields 92 percent citric acid with 8 percent soluble impurities.

Takeda developed a citric fermentation process utilizing paraffins as substrate. Pfizer used a paraffin-based process for several years too, when the cost of paraffin was relatively low. But raw material availability and cost no longer favor this route. Citric acid sells for about 40.85 cents/lb (anhydrous). The major citric acid producers include La Citrigue Belge, Tienen (Roche) in Belgium, Benckiser in Germany, Rhone-Poulenc in France, Jungbunzliner in Austria, and Miles and Pfizer in the United States. Up to 1967 there was a third U.S. producer, the original Stepan

Fermentation at Fieldsboro, New Jersey. The plant was sold to Bzura Chemical, and was subsequently closed down because of operating difficulties. Cargill became the third U.S. citric acid producers in 1989. Recently, Pfizer sold its citric operation to ADM. The largest Japanese producer is Owata Chemical, with an annual sale of citric acid at about 10 billion yen. The world production is estimated at 250,000 tons a year, with about half produced in the United States. The demand for citric acid is expected to grow continuously. To successfully produce citric acid and compete with the established producers, a manufacturer must have extensive process know-how.

Itaconic Acid

Itaconic acid (methylene succinic acid) is an unsaturated dibasic acid:

It is a structurally substituted methacrylic acid. Consequently, its principle use is as a copolymer in acrylic or methacrylic resins. Acrylic fibers, by definition, contain at least 85 percent acrylonitrile. Because pure acrylic fibers are dye-resistant, it is necessary to include other components to make the fibers susceptible to dyes. An acrylic resin containing 5 percent itaconic acid offers superior properties in taking and holding printing inks and in bonding. In addition to its main application as a component of acrylic fibers, itaconic acid is used in detergents, food ingredients, and food shortenings.

Previously, itaconic acid was isolated from pyrolytic products of citric acid or produced by converting aconitic acid present in sugar cane juice.

Fig. 24.19. Proposed metabolic sequence for biosynthesis of itaconic acid.

It now is produced on a commercial basis predominantly by direct fermentation of molasses.

The biosynthesis of itaconic acid once was believed to follow the decarboxylation of aconitic acid of the citric acid cycle. Now it is thought to follow the metabolic sequence shown in Fig. 24.19.

Although both *Aspergillus itaconicus* and *Aspergillus terreus* are known producers of itaconic acid, the latter is superior to the former and is believed to be used industrially. Either surface (shallow-pan) or submerged (deep-tank) fermentation can be used. The medium contains molasses, cornsteep liquor,

ammonium sulfate, and mineral salts of calcium, zinc, magnesium, and copper. The fermentation, similar to that of citric acid, is very sensitive to concentrations of copper and iron. Copper ion favorably restricts growth and product destruction, but excessive concentration of iron results in reduction of product accumulation. Ten to twenty percent (by volume) inoculum is used. The fermentation is carried out at around 40°C and a pH of 2.0 to 4.0. Vigorous agitation is employed. Moderate, but continuous, aeration is required. Air failure of very brief duration is enough to damage the fermentation. The batch cycle is 3 to 6 days. The highest known product concentration is 180 to 200 g/liter from a medium containing 30 percent sugar. The yield of itaconic acid on sugar is typically 50 to 70 percent. The itaconic acid recovery scheme involves the following:

1. Acidification of itaconic precipitates, if present.
2. Filtration to remove mycelium and other suspended solids.
3. Activated carbon treatment (this step and the next can be omitted for industrial-grade product).
4. Filtration to remove carbon.
5. Evaporation and crystallization.

If a high-purity acid is desired, further purification steps such as solvent extraction, ion exchange, and carbon decolorization are required.

The current price of itaconic acid is about $1/lb. The world production is estimated to be 10,000 tons a year. The major producers are Pfizer in the United States, Iwata in Japan, and Pfizer and Melle Bezous in Europe.

Gluconic Acid Aldehyde

Gluconic acid aldehyde is produced by the oxidation of the aldehyde group of glucose. Gluconic acid may be prepared from glucose by oxidation with a hypochlorite solution, by electrolysis of a solution of sugar containing a measured amount of bromine, or by fermentation of glucose by fungi or bacteria. The latter method now is preferred for economic reasons.

Gluconic acid is marketed in the form of 50 percent aqueous solution, calcium gluconate, sodium gluconate, and glucono-δ-lactone. Gluconic acid finds use in metal pickling, in foods as an acidulant, in tofu (soybean curd) manufacture as a protein coagulant, in detergent formulations as a calcium sequestrant, in the pharmaeutical area in mineral (calcium and iron) supplements, and in the construction area as a cement viscosity modifier. Calcium gluconate is widely used for oral and intravenous therapy. Sodium gluconate, a sequestering agent in neutral or alkaline solutions, finds use in the cleansing of glassware. Glucono-δ-lactone is used as a food flavor and an acidulant in baking powders and effervescent products.

The organism commonly used in gluconic acid fermentation is *Aspergillus niger* or *Gluconobacter suboxydans*. The larger-volume production uses the fungal process. Most of the *Gluconobacter* production is marketed as glucono-δ-lactone.

During gluconic acid fermentation, glucose is first oxidized (or, more correctly, dehydrogenated) to glucono-δ-lactone. This is carried out by glucose oxidase. Hydrogen peroxide also is produced in this step, but is decomposed by catalase.

The fermentation can be by either surface or submerged culture, the latter being the more generally practiced in industry. Horizontally rotating fermentors also have been used.

Calcium gluconate fermentation, in which calcium cabronate is used for neutralization of the product, is limited to an initial glucose concentration of approximately 15 percent because of the low solubility of calcium gluconate in water (4% at 30°C). The addition of borate or boric acid allows the use of up to 35 percent glucose in the medium. However, borogluconate was found deleterious to blood vessels of animals, and the product was withdrawn from the market.

The recovery of calcium gluconate from fermentation broth involves the following:

1. Filtration to remove mycelium and other suspended solids.
2. Carbon treatment for decolorization.
3. Filtration to remove carbon.
4. Evaporation to obtain a 15 to 20 percent calcium gluconate solution.
5. Crystallization at a temperature just above 0°C.

In sodium gluconate fermentation, sodium hydroxide is used to control pH. Sodium gluconate is much more soluble than calcium gluconate. The addition of sodium hydroxide provides an easy and precise means of neutralizing the acid as it is produced. Much higher concentration of glucose (up to 35%) can be used in this fermentation. The medium also contains cornsteep liquor, urea, magnesium sulfate, and some phosphates. The pH is controlled above 6.5 by addition of sodium hydroxide. One to 1.5 volumes of air per volume of solution per minute (vvm) are supplied for efficient oxygenation. High back pressure (up to 30 psig) is desirable. The fermentation cycle is 2 to 3 days. Continuous fermentation is used in Japan to convert 35 percent glucose solution to sodium gluconate with a yield higher than 95 percent. The continuous process doubles the productivity of the usual batch system.

Sodium gluconate can be recovered from fermentation filtrate by concentrating to 42 to 45 percent solids, adjusting to pH 7.5 with sodium hydroxide, and drum-drying.

In glucono-δ-lactone fermentation, *Gluconobacter suboxydans* converts a 10 percent glucose solution to glucono-δ-lactone and free gluconic acid in about 3 days. Approximately 40 percent of the gluconic acid is in the form of glucono-δ-lactone. Aqueous solutions of gluconic acid are in equilibrium with glucono-δ-lactone and glucono-γ-lactone. Crystals separating out of a supersaturated solution below 30°C will be predominantly free gluconic acid; from 30° to 70°C the crystals will be principally glucono-δ-lactone; and above 70°C they will be mainly the γ-lactone.

Prices of gluconates vary. Sodium gluconate (technical grade) is about \$1.30/kg, and calcium gluconate (USP grade) is \$2.70/kg. The actual selling prices sometimes have declined to as low as \$0.90/kg because of an increased and unbalanced supply. Dry sodium gluconate is the main form of gluconic acid/gluconate consumed in the United States. Liquid sodium acid gluconate and gluconic acid are used to a smaller extent. Gluconic acid is used in applications where the sodium part of sodium gluconate is not desired, such as in acid cleaners. The primary application for glucono-δ-lactone is as an acidulant in cottage cheese. The world production of gluconic acid is estimated to be around 25,000 tons a year. The major producers include Akzo America, Pfizer, PMP Fermentation products (acquired by Fujisawa in 1985), Focus Chemical (U.S.), Benckiser, Diosyth, Merck, Pfizer, Givandan (largely owned by Rhone-Poulenc), Roquette Freres (Europe), Fujisawa, and Kyowa Hakko (Japan).

2-Ketogluconic Acid

2-Ketogluconic acid may be produced by a bacterial fermentation involving various strains

of *Gluconobacter* or *Pseudomonas*. Selected strains of *Pseudomonas fluorescens* have been reported as giving the highest yield (up to 90%) when glucose and gluconate is used in the medium in highly aerated processes. Gluconic acid is an intermediate in the process.

2-Ketogluconic acid usually is recovered and shipped in the free acid state after centrifugation or filtration to remove cells. Calcium is removed by precipitation with sulfuric acid. The filtered acid may be shipped as a syrup or as a crystalline material after evaporation under reduced pressure and below 50°C.

Figure 24.20 illustrates the kinetic pattern for the fermentation of 10 percent glucose medium in the presence of an excess of calcium carbonate. A 20 percent glucose medium is used commercially.

The principal use of 2-ketogluconic acid is as an intermediate in the preparation of isoascorbic acid, now known to the trade as erythorbic acid. Erythorbic acid, its esters, and its salts are used as water- or fat-soluble

Fig. 24.20. Kinetics of 2-ketogluconic acid fermentation from glucose. (A: glucose; B: gluconic acid; C: bacteria; D: 2-ketogluconic acid.) (Microbial Technology, *2nd ed., Vol. 1, p. 384, Academic Press, New York, 1979.*)

antioxidants to retain color, flavor, and nutritive values in canned fruits and vegetables and in meats and meat products.

Fujisawa Pharmaceutical (Japan) is a major producer of 2-ketogluconic acid. Its yearly production is about 2,000 tons of 2-ketogluconic acid, with most of that exported to the United States, Canada, and European countries.

ORGANIC SOLVENTS

The organic chemicals that fall into this category and can be produced by fermentation include ethanol, butanol, acetone, 2,3-butanediol, and glycerol. 2.3-Butanediol and glycerol fermentations have been developed at laboratory and pilot-plant scales, but have not been commercialized. Ethanol, butanol, and acetone have been produced industrially by fermentation, but chemical synthesis is the manufacturing practice of choice for economic reasons. However, as price and availability of ethylene and propylene as feedstocks for the synthetic processes become subjects of concern, there is renewed interest in examining the fermentation pocesses as means of producing all or a portion of the future needs of ethanol, butanol, and acetone.

Ethanol

Ethanol is used in industrial solvents, thinners, detergents, toiletries, cosmetics, and pharmaceuticals, and, most important, as an intermediate for manufacturing other chemicals such as glycol ethers, ethyl chloride, amines, ethyl acetate, vinegar, and acetaldehyde. With the ever increasing price and dwindling supply of crude oil, ethanol fermented from grains and other renewable organic resources is in close competition with synthetic ethanol produced from ethylene.

In 1984, 150 million gallons of ethanol were produced synthetically in the United States, and around 500 million gallons of ethanol were produced by fermentation. Close to 1.1 billion gallons of fermentation alcohol will be produced in 1992. Synthetic ethanol is produced from ethylene via catalytic hydration. Its major U.S. producers include Shell Chemical, Union Carbide, National Distillers and Chemical, and Eastman Kodak. Fermentation ethanol is produced from sugar-containing materials such as grain products, fruits, molasses, whey, and sulfite waste liquor utilizing yeasts. ADM (Archer Daniels Midland Co.), Grain Processing, CPC International, A. E. Staley, and Publicker Industries are among the major producers of fermentation ethanol for industrial and fuel consumptions. ADM, with an annual production of approximately 600 million gallons, is the biggest of them all. Yeasts, particularly strains of *Saccharomyces cerevisiae*, are almost exclusively used in industrial ethanol fermentation. *Saccharomyces cerevisiae* tolerates ethanol concentrations up to about 20 percent (by volume), and has relatively fast fermentation rates. It converts over 85 percent of the available carbohydrates to ethanol and carbon dioxide under anaerobic conditions. Air or oxygen suppresses the formation of ethanol (the Pasteur effect). under aerobic conditions, a major portion of the carbohydrates goes to cell growth.

Ethanol is formed via glycolysis (the Embden–Meyerhof–Parnes pathway). The overall reaction starting from glucose can be written as follows:

$$C_6H_{12}O_6 \rightarrow 2C_2H_5OH + 2CO_2 + 31{,}200\ \text{cal}$$

The ethanol yield from the above equation is 51 percent by weight. Because carbohydrate is used also for cell growth and respiration, the overall yield of ethanol from total carbohydrate consumed is typically between 42 and 46 percent.

Ethanol fermentation can be conducted on nearly any carbohydrate-rich substrate. Molasses, which is the waste mother liquor remaining after the crystallization of sucrose in sugar mill operations, is widely used. Blackstrap molasses contains 35 to 40 percent sucroses and 15 to 20 percent invert sugars (glucose and fructose). High-test molasses contains 22 to 27 percent sucrose and 50 to 55 percent invert sugars. Most of the blackstrap molasses does not require the addition of other nutrients for ethanol fermentation. However high-test molasses requires considerable quantities of ammonium sulfate and other salts such as phosphates. The nonsugar

solids nutrient content of high-test molasses is about 7 percent, compared to 28 to 35 percent in blackstrap molasses.

In the molasses process, blackstrap or high-test molasses is charged into a mixing tank, where it is diluted with warm water to give a sugar concentration of 15 to 20 percent. Mineral acid is added to adjust the pH to between 4.0 and 5.0. The diluted and acidified molasses, called "mash," then is pasteurized, cooled, and charged into fermentor tanks, where about 5 percent yeast inoculum is added. The fermentation is carried out non-aseptically at 23 to 32°C. Antibiotics may be added to control possible contaminations. Because the overall reaction is exothermic, cooling is required. The fermentation takes 28 to 72 hours (averaging about 44 hours) to produce an ethanol concentration of 8 to 10 percent. Carbon dioxide normally is vented. If it is to be collected and recovered, the vent gas is scrubbed with water to remove entrained ethanol and then purified using activated carbon. The activated carbon is reactivated periodically using hot air or stream. The outgassing of the carbon dioxide from the fermentors provides sufficient agitation for small tanks. Mechanical agitation may be added for large fermentors. The fermentation may be conducted batchwise or continuously, with or without recycling yeast. Although continuous fermentation and/or cell recycle can significantly improve productivity and thus reduce required capital investment, it may have only a limited impact on lowering product costs. A major portion of the costs comes from raw materials.

After the fermentation is completed, the liquor, known as beer, is withdrawn from the fermentors, passed through heat exchangers, and pumped to the upper section of a beer still (or "whisky column"), where the alcohol and other volatiles such as aldehydes are distilled off as the overhead. The bottoms, known as slop or stillage, are processed into animal feeds, known as distillers' dried grains, which contain residual sugars, proteins, and vitamins.

The overhead from the beer column is passed through a heat exchanger and condensed. This condensate, known as high wines, contains 50 to 70 percent alcohol. It is charged into an aldehyde column (or "head column"), where aldehydes, esters, and other low-boiling impurities are separated as overhead. The stream from about the middle of the column is run into a refining column (or "rectifying column"). The tails from the aldehyde column and the weak fractions from the beer column, called low wines, are rerun with subsequent batches.

In the rectifying column, the heads containing a trace of aldehydes are returned to the aldehyde column. Near the top of the column, the azeotropic alcohol–water mixture of 95 percent alcohol is taken off, condensed, and run to storage. The higher-boiling alcohols, known as fusel oils, containing amyl, butyl, isobutyl, propyl and hexyl alcohols, are withdrawn farther down the column. The fusel oils amount to 0.5 percent of total carbohydrates consumed. Water is discharged from the bottom of the column. In order to prevent the diversion of industrial alcohol to potable uses, it is denatured by the addition of some material that renders the alcohol so treated unfit for use as a beverage. The 95 percent alcohol from the rectifying column is stored in government-bonded warehouses. The alcohol is either denatured, dehydrated, or sold (tax-free or tax-paid).

Anhydrous or absolute alcohol is produced by several methods. A third component such as benzene may be added, and the mixture distilled. The ternary azeotrope thus formed carries over the water, leaving behind anhydrous alcohol. Another method uses countercurrent extraction with a third component such as glycerine or ethylene glycol. The added component depresses the vapor pressure of the water and allows anhydrous alcohol to be distilled from the top of the extraction column. Both these methods are run using continuous columns.

A process flow diagram of ethanol fermentation using molasses as substrate is shown in Fig. 24.21.

Ethanol also can be produced by fermentation of starch, whey, and sulfite waste liquor. Grain fermentations require additional pretreatment

Fig. 24.21. Ethanol fermentation using molasses as a substrate. (Industrial Chemicals, *Faith, Keyes, and Clark, 4th ed., p. 357 (1975). Copyright © John Wiley & Sons, Inc., New York, and used by permission of the copyright owner.*

because yeast cannot metabolize starch directly. The grain (usually corn) is ground and heated in an aqueous slurry to gelatinize or solubilize the starch. Some starch-liquefying enzymes may be added in this step at lower temperatures. The liquefied starch then is cooled to about 65°C, and treated with malt (germinated, dried barley, which is rich in starch-hydrolyzing enzymes), or fungal amylase (produced by *Aspergillus niger*) to convert starch to oligosaccharides. Then yeast is added along with amyloglucosidase (or glucoamylase), which breaks down oligosaccharides into glucose. The subsequent fermentation and refining procedures are the same as those using molasses as raw materials.

Research to improve ethanol fermentation has focused on the development of continuous fermentation and improved yeast strains. Other related developments include processes for the continuous hydrolysis of wood coupled with continuous fermentation, and for the use of other biomass feedstocks, such as whey and cellulosic waste material. Developments in genetic engineering have made possible the development of new organisms (yeasts and bacteria) that can survive in higher concentrations of ethanol, tolerate higher sugar concentrations, grow at higher temperatures, and utilize starch or cellulose feedstock directly.

Gasohol

Gasohol, a fuel mixture of 10 percent alcohol and 90 percent gasoline, is a going business in the United States—and a growing one. It may become the most important use of fermentation ethanol.

Ethanol from grain fermentations has been made competitive as a liquid fuel in the United States because of technology improvements, especially in the area of efficient energy usage in production plants, and various government subsidies designed to reduce the nation's dependence on imported foreign oil. The incentives have included the following:

1. An extension through 1992 of the 5.4-cent federal tax exemption for every gallon of gasohol sold as fuel.
2. In more than 20 states, the elimination of state gasoline taxes for gasohol.
3. Tax credits, loans, and loan guarantees for biomass-based alcohol plant constructions.

These subsidies may become unnecessary in the future if the price of crude oil continues to inflate at a rate at least 5 percent higher than the prices of corn and other agricultural crops.

Fermentation ethanol also has the potential to become a major chemical feedstock. The

U.S. production of etylene was 29.2 billion pounds in 1979. It would take the equivalent of 3 billion bushels of corn to produce the same quantity of ethylene via ethanol fermentation and subsequent dehydration. The use of other substrates, such as wheat, sugar cane, sweet sorghum, whey, waste paper, and fast-growing trees, has to be explored.

To realize and expand the use of ethanol as chemical feedstock and liquid fuel, fermentation research and development on the following would be helpful:

1. Strains that tolerate higher ethanol concentrations.
2. Strains that ferment optimally at higher temperatures.
3. Strains capable of fixing CO_2 to give higher yields.
4. Strains capable of utilizing a wider range of carbohydrates.
5. Utilization of cellulosic or fibrous components of corn and other crops.
6. Fermentations under vacuum conditions.
7. Continuous fermentations with or without recycling of yeast.
8. Simultaneous and continuous product removal by physical or chemical means such as extraction or ultrafiltration.
9. More energy-efficient ways of recovering and dehydrating ethanol.

Butanol/Acetone

The original observation of butanol production by *Clostridia* was made by Pasteur, and acetone formation was noted by Schardinger. Interest in commercializing butanol–acetone fermentation occurred in 1909 primarily as a means of obtaining butadiene as raw material for synthetic rubber. In 1914 Weizmann established a working process to ferment starchy grains such as maize or corn to produce butanol, acetone, and ethanol using *Clostridium acetobutylicum*. With the outbreak of World War I, the production of acetone was of great interest for the manufacture of cordite. Large-scale operations were established in Canada, the United States, India, and elsewhere during the war period. Shortly after the war, Du Pont developed fast-drying nitrocellulose lacquers for the automobile industry, and butyl acetate was the solvent of choice for coating. Large quantities of butanol esters were needed as solvents, and butanol became the principal product of the butanol–acetone fermentation. By the 1930s, some butanol and acetone were being produced by chemical synthesis, and the butanol–acetone fermentation industry faced economic difficulties. This problem was solved by the discovery of the strains *Clostridium saccharo-acetobutylicum* and *Clostridium saccharo-butyl-acetonicum-liquefaciens*, which would ferment molasses, a cheaper raw material than starchy grains, and the industry thrived until the end of World War II. The major U.S. producers at that time were Commercial Solvents and Publicker Industries. They operated plants at Terre Haute, Indiana; Baltimore, Maryland; and Philadelphia, Pennsylvania. The butanol–acetone fermentations were conducted in large-scale equipment; fermentors of 50,000- to 500,000-gallon capacity commonly were used. At the present time, the petrochemical processes dominate, and most fermentation processes have been closed down. Fermentation of molasses to acetone and butanol, however, is probably still carried out by National Chemical Products in South Africa and Usina Victor SA in Brazil. Similar operations existed in the past in several countries and probably could be started up under appropriate economic circumstances.

In the grain fermentation process, 8 to 10 percent corn mashes were fermented (corn contains 70–72% starch on a dry basis). Fermentation yields were on the order of 29 to 32 grams mixed solvent per 100 grams starch used, with a solvent ratio of approximately 60–30–10 (butanol–acetone–ethanol, respectively). The organisms possessed good diastatic activity, so malting was not required. The cooked sterile cornmeal suspension was aseptically transferred to sterile fermentors, inoculated, and incubated for about 65 hours at 37°C, after which the solvents were recovered by distillation. The aqueous residue (slop or stillage) was concentrated in multiple-effect evaporators and drum-dried for use in animal and poultry feeds.

With the advent of molasses-fermenting strains, more rapid fermentations were attained (40–48 hours), and the solvents produced contained as much as 65 to 75 percent butanol, principally at the expense of ethanol. One hundred pounds of blackstrap molasses (containing 57 lb of sugar) was fermented into the following spectrum of products:

Product	*Yield*
Butanol	11.5 lb (solvent ratio = 68%) (yield on sugar = 20%)
Acetone	4.9 lb (solvent ratio = 29%) (yield on sugar = 9%)
Ethanol	0.5 lb (solvent ratio = 3%) (yield on sugar = 1%)
Total solvents	16.9 lb (yield on sugar = 30%)
Carbon dioxide	32.1 lb (yield on sugar = 56%)
Hydrogen	0.8 lb (yield on sugar = 1%)
Dried stillage	28.6 lb

The anaerobic *Clostridium* yields energy by converting glucose to acetyl-CoA, formate, CO_2, and H_2. The reducing power is then used to produce butanol and acetone via acetoacetyl-CoA. The overall reaction for butanol and acetone production can be pictured as follows:

$$2C_6H_{12}O_6 \rightarrow C_4H_9OH + CH_3COCH_3 + 5CO_2 + 4H_2$$

According to the equation the yields on sugar for butanol, acetone, carbon dioxide, and hydrogen are 21, 16, 60 and 2 percent, respectively.

Continuous fermentations were described by Russian workers. A plant in Dokshukin was operated in three batteries of seven to eight fermentors of 60,000 to 70,000 gallons. The continuous cycle was 40 to 90 hours, and the flow rate through the battery was 5000 to 10,000 gal/h. The feed contained 4 to 6 percent carbohydrates. A combination of raw material was used: molasses, flour, and hydrolysate containing pentoses. The continuous process gave a 20 percent productivity increase, and saved 142 lb of starch for every ton of solvents produced.

A considerable amount of care must be exercised in carrying out butanol–acetone fermentation. It is biologically unstable and may fail completely when contaminated. Numerous instances of contamination by bacteriophage were encountered commercially, and on several occasions plants had to suspend operation until the entire plant could be decontaminated. Accordingly, absolute clean-liness, experienced personnel familiar with phage symptoms, and the maintenance of a vigorous rapid fermentation are necessities in butanol–acetone fermentation.

The butaol–acetone–ethanol industrial fermentation of the 1950s made some remarkable advances in distillation efficiencies. Highlights included concentrating upon beer (fermentation mixture) stripping, early removal of water from solvents, thermo-compression on the beer still, continuous refining of butanol, and recovery of heat. A process flowsheet is shown in Fig. 24.22. When the fermentation is completed, the fermentor broth containing around 2 percent mixed solvents is pumped to a beer column where a 50 percent solvent mixture is taken off overhead, and distiller's slop is removed as bottoms. The slop may be dried and sold as animal feed. Another by-product is a mixture of carbon dioxide and hydrogen. The mixed-solvent vapors from the beer column are led to a batch fractionating column, from which three fractions (acetone, with B.P. 56.1°C, ethanol with B.P. 78.3°C, and butanol with B.P. 117.3°C) are removed overhead, leaving water as bottoms. The acetone and ethanol fractions are purified by conventional fractionation. The butanol fractions containing 70 percent butanol and 30 percent water are removed overhead. On condensation, two layers are formed. The top layer (80% butanol and 20% water) is returned to the butanol column, and the bottom layer (4% butanol and 96% water) is returned to the beer column. Approximately 35,000 lb stream is consumed for every ton of solvents produced.

The U.S. production of butanol and acetone was 650 million lb and 2.2 billion lb, respectively, in 1988. Butanol costs 26 cents/lb, and acetone 24 cents/lb. The fermentation industry has the potential to capture all or at least a portion of the market by concentrating

Fig. 24.22. Process flowsheet of butanol–acetone–ethanol fermentation. (Industrial Chemicals, *Faith, Keyes and Clark, 4th ed., p. 178. Copyright © John Wiley & Sons, New York, 1975. Used by permission of the copyright owner.*)

on research and development along the following lines:

1. Develop strains to tolerate high concentrations of butanol and acetone.
2. Develop homofermentive strains to improve product yields.
3. Utilize cheaper raw materials such as waste carbohydrates.
4. Develop continuous processes to improve productivities.
5. Improve fermentation stability through both strain and equipment enhancements.
6. Develop more energy-efficient separation methods such as extraction or reverse osmosis.

AMINO ACIDS

General

Amino acids in general can be represented by the following formula:

Because the amino group is on the α-carbon, the amino acids with this general formula are known as α amino acids. The α-carbon atom becomes asymmetric when R is not an H atom. Naturally occurring amino acids have an L-configuration. Amino acids are the building blocks of proteins, and the elementary composition of most proteins are similar; the approximate percentages are:

$$C = 50\text{–}55 \qquad N = 15\text{–}18$$
$$H = 6\text{–}8 \qquad S = 0\text{–}4$$
$$O = 20\text{–}23$$

Table 24.5 gives the structure of R, molecular weight, and elementary composition or each of the 20 amino acids commonly obtained on hydrolysis of proteins.

Amino acids can be obtained from purified proteins by chemical or enzymatic hydrolysis. They also can be isolated from industrial byproducts, extracted from plant or animal tissues, or synthesized by organic, enzymatic, or microbiological means. Amino acids that are produced industrially by fermentation include arginine, citrulline, glutamic acid, glutamine, histidine, isoleucine, leucine, lysine, ornithine, phenylalanine, proline, threonine, tryptophan, tyrosine, and valine.

Alanine and aspartic acid are produced commercially utilizing enzymes. In the case of alanine, the decarboxylation of aspartic acid by the aspartate decarboxylase of *Pseudomonas dacunhae* is commercialized. The annual world production of alanine is about 130 tons. Aspartic acid is produced commercially by condensing fumarate and ammonia using aspartase from *Escherchia coli*. This process has been made more convenient with enzyme immobilization technique. Aspartic acid is used primarily as a

TABLE 24.5 Twenty Common Amino Acids

			Elemental Composition (% wt)				
Amino Acids	*R–*	*M.W.*	*C*	*H*	*O*	*N*	*S*
Alanine	CH_3—	89	40	8	36	16	0
Arginine	H_2N—C(=NH)—NH—CH_2—CH_2—CH_2—	174	41	8	18	32	0
Aspargine	H_2N—C(=O)—CH_2—	132	36	6	36	21	0
Aspartic acid	HO—C(=O)—CH_2—	133	36	5	48	11	0
Cysteine	HS—CH_2—	121	30	6	26	12	26
Glutamic acid	HO—C(=O)—CH_2—CH_2—	147	41	6	44	10	0
Glutamine	H_2N—C(=O)—CH_2—CH_2—	146	41	7	33	19	0
Glycine	H—	75	32	7	43	19	0
Histidine	HC=C—CH_2— (N, C, NH, H)	155	46	6	21	27	0
Isoleucine	CH_3—CH_2—CH(CH_3)—	131	55	10	24	11	0
Leucine	CH_3—CH(CH_3)—CH_2—	131	55	10	24	11	0
Lysine	H_2N—CH_2—CH_2—CH_2—	146	49	10	22	19	0
Methionine	CH_3—S—CH_2—CH_2—	149	40	7	22	9	2
Phenylalanine	C$_6$H$_5$—CH_2—	165	66	7	19	8	0
Proline	CH_2—CH—COOH (CH_2, NH, CH_2)	115	52	8	27	12	0
Serine	$HOCH_2$—	105	34	7	46	13	0
Threonine	CH_3—CH(OH)—	119	40	8	40	12	0
Tryptophan	C$_8$H$_5$N—CH_2— (N, H)	204	65	6	16	14	0
Tyrosine	HO—C$_6$H$_4$—CH_2—	181	60	6	26	8	0
Valine	CH_3CH(CH_3)—	117	51	9	12	27	0

raw material with phenylalanine to produce aspartame, a noncaloric sweetener, tradenamed Nutrasweet. Production and sales of aspartame have increased rapidly since its introduction in 1981. World production of aspartic acid in 1988 was about 5000 tons. Tyrosine, valine, leucine, isoleucine, serine, threonine, arginine, glutamine, proline, histidine, citrulline, L-dopa, homoserine, ornithine, cysteine, tryptophan, and phenylalanine also can be produced by the enzymatic or fermentation method. Japanese companies such as Ajinomoto, Kyowa Hakko, and Tanabe Seiaka, and their foreign subsidiaries account for more than two-thirds

of world production of amino acids. The world production of amino acids is estimated at approximately 300,000 tons a year with a sales value of about 5 billion dollars.

Amino acids play important roles in many areas related to nutrition and medicine. Most amino acids are produced for medicinal purposes. Methionine, lysine, and tryptophan are three of the most important essential amino acids in the body. Both methionine and lysine are produced in large quantities (more than 10,000 tons a year) and have moderate prices ($5–10/kg). Methionine is synthesized chemically, and is used exclusively as a nutritional supplement. Lysine is mainly produced by fermentation, and is used largely as an animal feed additive. Tryptophan, synthesized chemically, has a relatively small market (20–50 ton a year) because of its high selling price (about $150/kg). Recombinant DNA technology has recently made commercial production of trytophan by fermentation.

Besides lysine and methionine, the only amino acid produced in large quantities is glutamic acid. It is produced exclusively by fermentation, with an estimated world production of 200,000 tons a year. Monosodium glutamate (MSG) is widely used for food seasoning.

The industrial fermentations of glutamic acid and lysine will be discussed in more detail.

Glutamic Acid

MSG was first produced using acid hydrolysis of wheat gluten or soybean protein in 1909 by Ajinomoto. In 1957, a glutamate-producing bacterium was isolated, and subsequent research and development brought about the economic production of glutamic acid by fermentation.

A large number of glutamic acid-producing microorganisms are known. A partial list is given below.

Bacteria	*Corynebacterium glutamicum* (Synonym *Micrococcus glutamicus*)
	Brevibacterium flavum
	Brevibacterium divaricatum
Fungi	*Aspergillus terreus*
	Ustilago maydis

Among them, *Corynebacterium glutamicum* is used most commonly in industry.

The fermentation medium contains a carbon source (glucose, molasses, or acetic acid), a nitrogen source (urea, ammonium sulfate, cornsteep liquor, or casein hydrolyzate), small amounts of mineral salts which supply potassium, phosphorus, magnesium, iron, and manganese, and a few (less than 5) micrograms of biotin per liter. The biotin requirement is the major controlling factor in the fermentation. When too much biotin is supplied for optimal growth, the organism produces lactic acid. Under conditions of suboptimal growth, glutamic acid is excreted. The metabolic pathway involved in the biosynthesis of glutamic acid from glucose is shown in Fig. 24.23. The lack or very low content of α-ketoglutarate dehydrogenase is a special characteristic of glutamic acid-producing microorganisms.

The fermentation is conducted aerobically in tanks with k_La (volumetric oxygen transfer coefficient) values around 300 millimoles O_2/liter/h/atm. If aeration is not adequate, lactic acid is produced, and the yield of glutamic acid is poor. Too much air is no good either; it gives even more lactic acid,

Fig. 24.23. Pathway for glutamic acid formation from glucose.

plus some α-ketoglutaric acid. The fermentation temperature is usually 28 to 33°C, and the pH optimal for growth and glutamic acid production is 7.0 to 8.0. Continuous feeding of liquid ammonium hydroxide or gaseous ammonia controls pH and supplies ammonium ions to the fermentation. Modern digital computers are in some cases used for event sequencing, control, and optimization of the fermentation operation.

The fermentation cycle is 24 to 48 hours. The final concentration of glutamic acid is about 150 g/liter if glucose is used, and about 120 g/liter if molasses is used. The overall yield of glutamic acid on sugar is about 65 percent on a weight basis. A portion of sugar is used for cellular growth with about 50 percent yield, whereas the sugar used for glutamic acid production has a yield of 86 percent according to the following equation:

$$C_{12}H_{22}O_{11} + 2NH_3 + 3O_2 \rightarrow 2C_5H_9O_4N + 2CO_2 + 5H_2O$$

The recovery of glutamic acid from fermentation broth is relatively easy. The broth is clarified by adding acid to pH 3.4, heating to 87°C, holding for a sufficient time to coagulate suspended solids, and filtering the coagulated broth to yield a clarified broth. The clarified broth is concentrated by evaporation, and then crystallized to yield glutamic acid. Other recovery schemes exist.

MSG is an important flavor enhancer for natural and processed foods. It is also good for protecting the flavor and color of preserved foods and suppressing off-flavors. MSG is selling for about $2/kg, and glutamic acid (99.5%) about $4/kg. Glutamic acid is not an essential amino acid. However, it has some pharmaceutical uses. It also improves the growth of pigs.

The major manufacturers of glutamic acid include Ajinomoto, Asahi Foods, Kyowa Hakko, Takeda (Japan), Orsan (France), Biacor (Italy), Cheil Sugar, Mi-Won (S. Korea), and Tung Hai Fermentation Industry, Ve Wong (Taiwan). U.S. producers Merck, IMC, Great Western Sugar, and Stauffer ceased their MSG production in 1970, 1975, 1977, and 1984, respectively. To reduce production costs, most of the Japanese production is being moved to Southeast Asia or South America where raw material is available locally, in abundance and at lower cost—a trend for the manufacturing of all low-margin and high-volume fermentation products.

Lysine

Lysine, biologically active in its L-configuration, is an essential amino acid in human and animal nutrition. It is contained in good measure in fishmeal and in soybean meal. Plant products, particularly cereal grains such as corn, wheat, and rice, are usually low in lysine. Because of lysine deficiency, these grain proteins are of poorer quality than animal-derived proteins that contain higher levels of lysine and command much higher prices. For applications in human food, lysine in its salt forms can be added to cookies and bread, and in solution it can be used to soak rice. As most animal feed rations are based on maize and other grains, supplementing the feed-stuffs with lysine (plus methionine) significantly improves their nutritional value for breeding poultry and pigs.

Lysine can be obtained by several different processes:

1. Isolation from natural sources.
2. Chemical synthesis/enzymatic racemization.
3. Enzymatic conversion.
4. Fermentation.

Chemical pathway to lysine, although perfectly practical and giving high yields, results in a DL-racemic mixture. The resolution of the mixture to the L-configuration is costly.

The enzymatic process uses DL-aminolactam as starting material. It consists of the racemization of D-aminolactam and the hydrolyzation of L-aminolactam to form lysine. The reactions are shown in Fig. 24.24. One hundred g/liter of DL-aminolactam can be converted into lysine in 25 hours with 100 percent molar yield.

Fig. 24.24. Enzymatic steps to form lysine from DL-aminolactam.

Fig. 24.25. Two-stage fermentation for lysine production.

The fermentation process to produce lysine was first developed by Pfizer. *Escherichia coli* synthesized its own lysine requirements by converting carbohydrate and ammonia to α,ε-diaminopimelic acid (DAP). Its decarboxylation resulted in lysine. The industrial fermentation employed an *E. coli* mutant devoid of DAP-decarboxylase. It accumulated substantial DAP in the culture medium. After maximum DAP yields were attained, a second organism, another *E. coli* strain or *Aerobacter aerogenis*, was used as a source of DAP-decarboxylase to produce lysine. This second organism was devoid of lysine decarboxylase. Lysine, thus, accumulated in the broth. The reactions are shown in Fig. 24.25.

A single-stage or directed fermentation process for producing lysine later was made possible by using mutants of *Micrococcus glutamicus (Corynebacterium glutamicum)* or *Brevibacterium flavum.*

Both auxotrophic and regulatory mutants have been obtained for overproduction of lysine. Figure 24.26 shows the biosynthetic pathway of lysine and its metabolic controls in *Corynebacterium glutamicum.* The key enzyme, aspartate kinase, is under the influence of the concerted feedback inhibition by lysine and threonine. Threonine, alone, also inhibits the activity of homoserine dehydrogenase. Excess homoserine increases the threonine pool within the cells, and aspartate kinase then is inhibited. An auxotrophic mutant that requires homoserine (or threonine plus methionine) and lacks homoseine dehydrogenase, when grown under suboptimal homoserine concentrations, alleviates the feedback inhibition on aspartate kinase and shunts ASA to form lysine.

In the case of *Brevibacterium flavum*, only aspartate kinase is sensitive to the feedback inhibition.

Fig. 24.26. Regulation in lysine biosynthesis in *Brevibacterium flavum* and *Corynebacterium glutamicum*. (ASA: aspartate semialdehyde; DHDP: dihydrodipicolinate; Hse: hemoserine; DAP: diaminopimelate.) (Microbial Technology, *2nd ed., Vol. 1, p. 220, Academic Press, New York, 1979.*)

Regulatory mutants, whose aspartate kinase is desensitized to the concerted feedback inhibition, have been obtained by isolation of lysine analogue–resistant mutants, whose growth was not inhibited by a lysine analogue, such as amino ethyl cysteine (AEC). The lysine analogue behaved as a false feedback inhibitor on aspartate kinase and did not allow growth of the parent strain. Only the mutants desensitized to the feedback inhibition grew.

The *Corynebacterium glutamicum* mutant generally is used industrially in direct fermentation of lysine. Molasses is the most common carbon source. Sufficient amounts (over 30 micrograms/liter) of biotin must be included in the medium to prevent the excretion of glutamic acid. This biotin requirement usually is met by using molasses as the carbon source. The fermentation runs at temperature about 28 to 33°C, and pH 6 to 8. High aeration is desirable. The final product concentration is around 60 g/liter, and the fermentation cycle is 48 to 72 hours. The yield of lysine on carbohydrate is about 40 percent. The formation of lysine from molasses can be represented as follows:

$$C_{12}H_{22}O_{11} + 5O_2 + 2NH_3 \rightarrow C_6H_{14}O_2N_2 + 6CO_2 + 7H_2O$$

Ion-exchange resins are used for isolation of lysine from fermentation broths. The eluted lysine then is crystallized from water. The most common commercial form of lysine used in animal feed is 98 percent lysine monohydrochloride. It is equivalent to 78.4 percent of the amino acid lysine, which can be metabolized by animals to body proteins. The supplementation level is about 0.5 percent lysine in feed. When the *Brevibacterium* mutant is used to produce lysine, the entire fermentation broth may be evaporated and dried, and the dried product used as animal feed supplement.

The standard commercial food-grade lysine (as 98.5% lysine monohydrochloride) is about \$7/kg, and lysine as feed supplement is about \$3/kg. Its price varies and ties in closely with the cost and availability of fishmeal and soybean cake. The annual world production of lysine is about 80,000 tons with Ajinomoto, Kyowa Hakko, Toray (Japan), Mi-Won (S. Jorea), Rhone Poulenc, Orsan (France), and Biokyowa, Heartland Lysine (U.S.) being the major producers.

Ajinomoto is expanding its lysine production capacity at Heartland Lysine in the United States to 12,000 tons/year, Thai Ajinomoto in Thailand to 6,000 tons and at Eurolysine to 40,000 tons. Kyowa Hakko is expanding its production capacity at Biokyowa in

the United States to 13,000 tons/year and Fermentacions Mexicanas in Mexico to 10,000 tons, and just completed a 5000 tons/year production facility at Agroferm in Hungary.

The world demand for lysine is expected to increase continually. The increasing ethanol fermentation capacity for gasohol production provides an opportunity for increased lysine production. Distillers' dried grains (DDG), a by-product of ethanol fermentation used as a protein source in animal feeds, is deficient in lysine and other essential amino acids. It needs to be supplemented with lysine for full-value use.

VITAMINS

Many vitamins of medical importance can be synthesized by microorganisms. However, only "two and a half" vitamins are now produced by fermentation. Vitamin B_2 (riboflavin) and vitamin B_{12} (cyanocobalamin) are products of fermentation. The remaining half vitamin produced by fermentation is vitamin C. L-Sorbose is the precursor of vitamin C (ascorbic acid), and L-sorbose is produced mirobiologically from sorbitol.

Vitamin C (Ascorbic Acid)

Figure 24.27 gives the reactions involved in the combined microbiological and chemical conversions of glucose to vitamin C. Sorbitol is made by catalytic hydrogenation of glucose. L-Sorbose is produced from sorbitol by the action of several species of bacteria of the genus *Acetobacter*. The most commonly used microorganism is *Acetobacter suboxydans*. Because this organism is very sensitive to nickel ions, it is important that the medium and fermentor be free of nickel. The medium normally consists of 100 to 200 g/liter sorbitol, 2.5 g/liter cornsteep liquor, and an antifoam such as soybean oil. The medium is sterilized and cooled to 30 to 35°C, and about 2.5 percent inoculum is added. The tank is aerated and sometimes stirred. Yields of 80 to 90 percent of the sugar used are commonly obtained in 20 to 30 hours.

The chemical steps in the conversion of sorbose to ascorbic acid involve the preparation of the diacetone derivative, which then is oxidized; the acetone groups are removed, and the resultant 2-keto-L-gulonic acid is isomerized to the enediol with ring closure.

Sorbitol, sorbose, and ascorbic acid are selling for about \$1.30, \$3, and \$10/kg, respectively. The estimated world production of sorbose is around 25,000 tons a year, with about half of it produced in the United States.

Alternative vitamin C production technologies have been pursued by groups all around the world. One-step fermentation starting from glucose to produce 2-keto-L-gluconic acid was reported to be practiced in China, a microalgae process was developed in the U.S. and a two-stage fermentation via the intermediate 2,5-diketo-D-guconic acid was developed in Japan. A U.S. group, on the other hand, succeeded in cloning the 2,5-diketo-D-gluconate reductase gene from a *Corynebacter sp.* and expressing it in an *Erwina sp.* which naturally produces 2,5-diketo-D-gluconic acid via a glucose oxidation fermentation. The recombinant *Erwina* culture is thus capable of carrying out the complex oxidation and reduction steps to form 2-keto-L-gluconic acid in a single fermentation.

Fig. 24.27. Reactions leading to vitamin C from glucose.

Vitamin B_2 (Riboflavin)

Microbiologically produced riboflavin (structure shown in Fig. 24.28) has long been available in yeast and related preparations in association with many other vitamins of the B-complex. Aside from yeast, the first organism employed primarily for riboflavin production was *Clostridium acetobutylicum*, the anaerobic bacterium used for the microbial production of acetone and butanol. Riboflavin was purely a by-product and was found in the dried stillage residues in amounts ranging from 40 to 70 micrograms per gram of dried fermentation solids. Further research developed improvements, adaptable only to the fermentation of cereal, grains, and milk products by *Clostridium acetobutylicum* to yield residue containing as much as 7000 micrograms of vitamin B_2 per gram of dried solids. This was effected principally by reducing the iron content of the medium to 1 to 3 ppm, and fermenting in stainless steel, aluminium, or other iron-free tanks.

Later investigations disclosed that riboflavin could be produced by species of a yeast, *Candida flareri* or *C. guilliermondi*, when grown under aerobic conditions in a medium containing a fermentable sugar, an assimilable source of nitrogen, biotin, and less than 100 micrograms of iron per liter of medium. Yields as high as 200 mg/liter were obtained.

Other studies on a fungus, *Eremothecium ashbyii*, and a closely related organism known as *Ashbya gossyii* resulted in the production of much larger amounts of riboflavin. An aerobic process was used in which the iron content was not critical. Riboflavin was produced in large amounts by the fermentation industry using either the *Eremothecium* or *Ashbya* strains. Yields as high as 10 to 15 g/liter were possible.

Fig. 24.28. Riboflavin: vitamin B_2.

Late in the 1960s, concurrently with the development of the riboflavin fermentation process, a synthetic means of producing riboflavin was discovered. This synthetic process dominated the production of riboflavin until 1972 when some major fermentation strain and process improvements were made with the *Ashbya gossypii* strain. Since then yields have been significantly improved. The fermentation method now accounts for essentially all the riboflavin produced.

The fermentation lasts 8 to 10 days. Cell growth occurs in the first 2 days, and enzymes catalyzing riboflavin synthesis are formed during the growth period. Glycine and edible oil stimulate the formation of riboflavin, but they are not its precursors. The additions of carbohydrate and oil permit the overproduction of riboflavin. The C/C yield is about 50 percent on carbohydrate, and about 100 percent on oil.

Upon completion of the fermentation, the solids are dried to a crude product for animal feed supplement or processed to a USP-grade product. In either case, the pH value of the fermented medium is adjusted to pH 4.5. For the feed-grade product, the broth is concentrated to about 30 percent solids and dried on double-drum driers.

When a crystalline product is required, the fermented broth is heated for one hour at 121°C to solubilize the riboflavin. Insoluble matter is removed by centrifugation, and riboflavin recovered by conversion to the less soluble form. Both chemical and microbiological methods of conversion have been used. The precipitated riboflavin then is dissolved in water, polar solvents, or an alkaline solution, oxidized by aeration, and recovered by recrystallization from the aqueous or polar solvent solution or by acidification of the alkaline solution.

The major producers of riboflavin include Merck, Hoffmann-LaRoche, and Pfizer in the United States with an estimated annual production of 1.3-million kilograms. The feedgrade product sells for about $35 a kilogram and the USP-grade product $55 a kilogram.

Vitamin B_{12} (Cyanocobalamin)

Vitamin B_{12}, cyanocobalamin, is an important biologically active compound. It serves as a hematopoietic factor in mammals and as a growth factor for many microbial and animal species. Its markets are divided into pharmaceutical (96–98% pure) and animal feed (80% pure) applications. All vitamin B_{12} now is made commercially by fermentation.

The microbiological production of vitamin B_{12} arose from an interesting sequence of events. For many years, liver extract was used to check cases of pernicious anemia. Investigators at Merck discovered that crystalline extracts made from liver tissue contained the highly active compound responsible for the therapeutic action. Identity with the antianemia factor in liver then was established, and the compound was called vitamin B_{12}. Later it was found that spent liquors from streptomycin and other antibiotic fermentations contained appreciable amounts of vitamin B_{12}. Soon vitamin B_{12} derived from cultures of these fermentations supplanted beef liver as a practical source of the vitamin. Around 1950, materials rich in biomass such as *Actinomycetes* or bacteria broths of antibiotic fermentations or dried sewage residues of activated sludge processes were used for isolating vitamin B_{12}, either in a crude form for animal feeding or in a pure state for medicinal uses. Later, high-producing bacterial strains were specially selected for commercial production. Today vitamin B_{12} is obtained from fermentations using selected strains of *Propionibacterium* or *Pseudomonas* cultures. A full chemical synthesis of vitamin B_{12} is known. However, it requires some 70 steps and is of little value for all practical purposes.

The *Pseudomonas denitrificans* strain is most often used for commercial production of vitamin B_{12}. It only requires traditional components in the growth medium, such as sucrose, yeast extract, and several metallic salts. Dimethylbenzimidazole (10–25 mg/liter) and cobaltous nitrate (40–200 mg/liter) must be supplemented at the start of the culture in order to enhance vitamin production. Betaine stimulates the biosynthesis of vitamin B_{12}, even though it need not be metabolized by the organism. Choline also has favorable effects in activating some biosynthesis steps or altering the membrane permeability. Glutamic acid, on the other hand, stimulates cellular growth. Because of its relative cheapness and high betaine and glutamic acid contents, beet molasses (60–120 g/liter) is preferentially used in industrial fermentations of vitamin B_{12}. The fermentation is conducted with aeration and agitation. Temperature is optimal around 28°C, and pH optimal around 7.0. The yield reported in the literature was 59 mg/liter in 1971, using a *Pseudomonas* strain. A yield of 200 mg/liter was reported for vitamin B_{12} fermentations using *Propionibacteria* in 1974. It is believed that yields of vitamin B_{12} have been significantly improved since then.

The isolation of vitamin B_{12} from fermentation broth, where it is normally present in parts per million, is a brilliant achievement on the part of the chemist and chemical engineer. About 80 percent of the vitamin produced is outside the cells, and 20 percent inside the cells. The whole broth is heated at 80 to 120°C for 10 to 30 minutes at pH 6.5 to 8.5. The heated broth is treated with cyanide or thiocynate to obtain cyanocobalamin. The separation then can be accomplished by adsorption on a cation-exchange resin, such as Amberlite IRC 50. Extraction also can be done by using phenol or cresol alone or in a mixture with benzene, butanol, carbon tetrachloride, or chloroform; or it can be done by precipitation or crystallization upon evaporation with appropriate diluents such as cresol or tannic acid. Using the extraction method, 98 percent pure cyanocobalamin can be obtained with a 75 percent yield.

The total world market for cyanocobalamin is estimated about 600 kilograms a year. The USP-grade vitamin B_{12} is selling for \$10–15 a gram. The major producers include Merck (U.S.), Rhone-Poulenc, Roussel (France), Glaxo (England), Farmitalia (Italy), Chinoin, and Richter (Hungary).

ENZYMES

All fermentation processes are the result of the enzyme activity of microorganisms. In fact, life itself, whether plant or animal, involves a complex myriad of enzyme reactions.

An enzyme is a protein that is synthesized in a living cell. It catalyzes or speeds up a thermodynamically possible reaction so that the rate of the reaction is compatible with the numerous biochemical processes essential for the growth and maintenance of a cell. An enzyme, like chemical catalysts, in no way modifies the equilibrium constant or the free energy change of a reaction. The synthesis of an enzyme thus is under tight metabolic regulations and controls sometimes that can be genetically or environmentally manipulated to cause the over-production of an ezyme by the cell.

Being a protein, an enzyme loses its catalytic properties when subjected to agents such as heat, strong acids or bases, organic solvents, or other materials that denature the protein. Each enzyme catalyzes a specific reaction or a group of reactions with certain common characteristics. The high specificity of the catalytic function of an enzyme is due to its protein nature; that is, the highly complex structure of a protein can provide both the environment for a particular reaction mechanism and the template function to recognize a limited set of substrates.

Enzymes are used quite extensively now as industrial catalysts. They offer the following advantages in comparison with chemical catalysts:

1. They are specific in action, and thus minimize the occurrence of undesirable side reactions.
2. They are relatively cheap when used in crude form.
3. They are effective for chemical conversions within the physiological range of pH and at low temperatures and pressures.
4. They are relatively nontoxic and thus are acceptable for applications in food processes and medicinal treatments.
5. They are effective within a wide range of substrate concentrations.

The enzymes of most economical importance, such as amylases, glucose isomerase, and proteases, will be discussed individually. Altogether they account for almost 90 percent of the total sales of enzymes. Other industrially important enzymes with relative small sales volumes are listed in Table 24.6, along with their microbial sources and commercial applications. In addition, there are many microbial enzymes that are used for analytical, clinical, and research purposes; they include hexokinases, pyruvate kinase, uricase, glucose-6-phosphate dehydrogenase, amino acid oxidase, aminopeptidase, and others. Restriction enzymes, endonucleases, have been used widely in recombinant DNA research. Glycosyltransferases have received much attention recently as glycobiology research picks up steam.

The world sales of enzymes is estimated at about \$600 million a year. About 60 percent of the enzymes are proteolytic, for use by the detergent, dairy, and leather industries. The carbohydrates, mainly the amylases, isomerases, and pectinases, used in baking, brewing, distilling, starch, and textile industries, represent nearly 30 percent of the total enzyme market. Lipase and other highly specialized enzymes make up the remaining 10 percent of the total enzyme usage. Novo Industri (Denmark) and Gist-Brocades (Holland) have the largest share of the enzyme market. A large portion of U.S. production of industrial enzymes is for "captive" uses by the producing companies.

Amylases

The commercial importance of amylolytic enzymes is rapidly increasing. These enzymes catalyze the hydrolytic reactions of amylose (the unbranched starch) and amylopectin (the branched starch). Amylases, according to

TABLE 24.6 Some Industrially Important Enzymes Other Than Amylases, Proteases, and Glucose Isomerase

Enzyme	*Microbial Sources*	*Applications*
Amino acylase	Bacteria	*L-Amino acid production.
Asparginase	*Escherichia coli*	*Treatment of acute lymphatic anaemia. *Remission of lymphatic leukemia. *Anticancer therapy.
Catalase	*Aspergillus niger*	*Removal of trace of hydrogen peroxide. *Used with glucose oxidase.
Cellulase and hemicellulase	*Trichoderma reesei* *Aspergillus niger*	*As a digestive aid. *Production of sugar syrup.
Dextranase	*Penicillium funiculosum* *Trichoderma sp.*	*Prevention of dental plaque. *Removal of dextran impurities in sugar refining.
Glucose oxidase (Notatin)	*Aspergillus niger* *Penicillium notatum*	*As oxygen scavenger in food industry. *Combined with catalase for removal of glucose or oxygen. *Combined with peroxidase for quantitative determination of glucose.
Invertase	*Saccharomyces cerevisiase*	*Sucrose hydrolysis to form glucose and fructose which are sweeter and have lower crystallinity. *Used in jam making and chocolate manufacturing.
Lactase	*Kluyveromyces fragilis* *Kluyveromyces lactis* *Lactobacillus sp.* *Saccharomyces lactis*	*Lactose hydrolysis to form glucose and galactose which are sweeter and more soluble.
Lipase	*Aspergillus niger* *Candida lipolytica* *Mucor javanicus* *Rhizopus arrhizus*	*As a digestive aid. *Used in waste treatment. *Used in tanning. *Used to improve flavor.
Pectinase	*Aspergillus niger* *Aspergillus sp.*	*Clarification of fruit juice. *Increased juice extraction. *Pectin decomposition and viscosity reduction.
Penicillin acylase (Penicillin amidase)	*Bacillus megaterium* *Escherichia coli*	*Formation of 6-APA from penicillin for production of semisynthetic pencillins.
Pullulanase	*Aerobacter aerogenes* *Klebsiella aerogenes*	*Increase of fermentability of starch worts or syrups. *Structural determination of polysaccharides.
Rennets	*Bacillus polymyxa* *Bacillus subtilis* *Endothia parasitica* *Mucor miehei* *Mucor pusillus*	*As coagulants in cheese curd making.

their difference in modes of action, can be divided into:

1. α-Amylase, which hydrolyzes α-1,4-linkages randomly to yield a mixture of oligosacharides, maltose, and glucose.
2. β-Amylase, which cleaves away successive maltose units from the nonreducing end of starch to yield maltose quantitatively.
3. Glucoamylase (or amyloglucosidase), which chops off glucose successively to yield glucose.
4. The debranching enzymes such as pullulanase, which attacks the α-1,6-linkages at the braching point of amylopectin.

Amylases are used extensively in the following ways:

- For production of grain syrup, glucose syrup, liquid glucose, and crystalline glucose.
- For production of high-fructose corn syrup in connection with glucose isomerase.

- To solubilize and saccharify starch for alcohol fermentation in brewing, distilling, and fuel industries.
- To modify the viscosity of starch used in coating printing papers.
- To remove starch sizes applied to cotton thread before weaving in the textile industry.
- For production of maltose-containing syrups in brewing and baking industries.
- To reduce the viscosity of sugar syrups used in various food and sugar products.
- As a component in digestive aids.

α-Amylase is produced commercially by using both fungal and bacterial species. The fungal amylase has a relatively low heat stability, and its major application is in the baking industry to supplement the variable activity of the amylase present in wheat flour. The bacterial amylase is much more heat-stable, and it is used in brewing, starch degradation, alcohol, and textile industries. The organisms commonly used for the commercial production of α-amylase include:

Fungi	*Aspergillus oryzae*
Bacteria	*Bacillus subtilis*
	Bacillus licheniformis
	Bacillus amyloliquefaciens

Aspergillus oryzae (the green fungus) can be grown in either semisolid or submerged culture. In semisolid culture it produces several enzymes, primarily α-amylase, glucoamylase, lactase, and protease. In submerged culture the production of α-amylase is increased, and the formation of other enzymes becomes minimal. The use of this fungal amylase in the baking industry speeds up the yeast (*Saccharomyces carlsbergenis*) fermentation, produces stiffer, more stable doughs, and improves the texture, porosity, digestibility, and shelf life of bread. The fungal α-amylase delivers its optimal activity at 5 to 7 pH and at 50 to 55°C.

Different amylase-producing organisms may require different fermentation conditions for optimal enzyme production. When *Bacillus subtilis* is used, the fermentation medium may contain starch, cornsteep liquor, yeast, phosphate, and some mineral salts. The amylase treatment on starch is often short to prevent the significant accumulation of glucose, which is inhibitory to the *Bacillus* amylase fermentation. The fermentation is run at neutral pH and at around 35°C. Care must be taken to prevent contamination. The time cycle is about 48 hours. Whole mash may be used directly for starch liquefaction and saccharification, or the mash may be processed to produce liquid or crystal enzyme preparation with high purity. The processing, involving filtration or centrifugation of the bacterial fermentation broth, presents real problems to the recovery plant. Pretreatment with a coagulating or flocculating agents often is needed. The amylase produced by this *Bacillus* strain is relatively unstable, but the addition of calcium chloride improves the stability.

Following the solubilization of starch by α-amylase (often of a bacterial origin), further degradation is achieved using fungal glucoamylase. *Aspergillus niger* (the black fungus) is commonly used for the production of glucoamylase. The fungal fermentation starts with a medium containing 25 to 30 percent starch and around 10 percent cornsteep liquor. Incremental or continuous feeding of concentrated nutrients may be used to circumvent the problems caused by a concentrated initial medium. The fermentation pH is about 4.0 and temperature around 28°C. The fermentation has a high oxygen demand. High oxygen tension, however, inhibits enzyme production. Zero D.O. (dissolved oxygen) is not atypical in this fermentation. After the completion of the fermentation in 4 to 5 days, the fermentor mash is cooled and filtered to remove cells and insoluble matters. Transglucosidase may be removed using clay, destroyed preferentially using proteases at certain pH's and temperatures, or inactivated by magnesium oxide. Contamination of glucoamylase by the transglucosidase activity may result in loss of 5 to 10 percent of glucose to form isomaltose and panose by a reversion process. These reversion products also impede the crystallization of glucose.

Fungal glucoamylase in combination with bacterial α-amylase make a complete ezymatic mixture for hydrolysis of starch to glucose. Prior to liquefaction, starch is gelatinized by heat treatment at temperatures above 100°C. The liquefaction to form maltodextrins is aided by the action of bacterial α-amylase. α-Amylase from *Bacillus amyloliquefaciens* functions at 5.5 to 7 pH and 90°C, and α-amylase from *Bacillus licheniformis* functions at 5.5 to 9 pH and temperatures as high as 110°C. In the subsequent saccharification process, an appropriate amount of glucoamylase from *Aspergillus niger* is added to the thinned starch (30–50% dry substance) with stirring at 55 to 60°C and 4 to 5 pH for 48 to 72 hours. This achieves a final D.E. (dextrose equivalence) of about 97, with about 94% of the dry weight being glucose. The equilibrium concentrations of the saccharides formed by resynthesis limit the maximum degree of hydrolysis obtainable. Because the activity of glucoamylase toward the branching points (the α-1,6-linkages) is low, it may be advantageous to use a debranching enzyme such as pullulanase early in the hydrolysis process.

The enzymatic hydrolysis of starch to glucose is commercially preferred to the acid hydrolysis route using hydrochloric acid. The enzymatic process produces fewer side products, does not involves a corrosive acid, and allows the use of less pure starch products whose protein contaminants would, upon acid hydrolysis, give amino acids and browning reactions.

Glucose Isomerase and HFCS (High Fructose Corn Syrup)

Starch degradation using α-amylase and glucoamylase produces corn syrup with glucose concentration up to 94 percent on a dry weight basis. The glucose then can be isomerized to a mixture of glucose and fructose by glucose isomerase, which is present in many microorganisms. The product, commonly known as "high fructose corn syrup" (HFCS), typically contains 42 percent fructose, 50 percent glucose, 6 percent maltose, and 2 percent maltotriose. HFCS is also available in the form of 55 percent syrup, 90 percent syrup, or 99 percent crystal. Figure 24.29 gives a block diagram that shows how a corn wet-milling processor produces its starch, corn syrup, glucose, and fructose syrups.

Fructose, the monosaccharide commonly called fruit sugar, is about 50 percent sweeter than sucrose, the disaccharide familiarly known as table sugar. Sucrose can be hydrolyzed to invert sugar which is a mixture of fructose and glucose. Liquid invert sugar has been the major sweetener used in soft drinks and as a food ingredient.

The 42 percent HFCS sells for about 20 cents/lb, which is 15 to 20 percent cheaper than liquid invert sugar on a dry weight basis. The soft-drink industry, shifting away from liquid invert sugar, is the major user of 42 percent HFCS. The baking industry ranks as the second largest user.

The 55 percent HFCS sells for about 25 cents/lb, which is 5 to 10 percent cheaper than liquid invert sugar on a dry weight basis. Fifty-five percent HFCS, made by combining the 42 percent HFCS with 90 percent HFCS, has about the same degree of sweetness as sucrose. It is used as a sweetener and flavor-enhancer in fruit-flavored soft drinks. Fructose enhances flavors, whereas sucrose masks them.

Fig. 24.29. Flowsheet of a corn wet-milling process.

The 90 percent HFCS, 40 to 50 percent sweeter than sucrose, sells for about 45 cents/lb, which is 40 to 50 percent more expensive than liquid invert sugar on a dry weight basis. Because it is much smaller than sucrose, a smaller amount of it is needed to sweeten a product to desired levels, and sweetener calories in the product can be reduced by about one-third. The major use of 90 percent HFCS is in dietetic foods and drinks.

The 99 percent fructose in crystal form is obtained by drying the 90 percent HFCS. It is about 70 percent sweeter than sucrose. Being an essentially pure sweetener, it allows the full taste of product flavors to develop. It is finding uses in diabetic and low calorie foods and drinks.

The high fructose corn syrup world production totals 2 million tons a year. The glucose isomerase business amounts to about $50 million a year. Among the major HFCS producers are ADM, Car-Mi (a joint venture of Cargill and Miles), A. E. Staley, Clinton/Standard Brands, CPC, and Hubinger/Heinz.

Many organisms are glucose isomerase producers. Most of them produce xylose isomerase with low specificity, and glucose can be an alternative substrate for the enzyme. Table 24.7 lists microorganisms believed to be used industrially for glucose isomerase production, along with their reported yields.

The desirable application conditions of glucose isomerase are 7.0 to 8.0 in pH and around 65°C in temperature.

Glucose isomerase fermentation typically has a cycle time of about 2 days. The fermentation conditions vary from producer to producer. Many glucose isomerase–producing organisms require xylose and cobalt for full enzyme induction. Xylose is too rare and expensive to be used in commercial fermentation processes, and cobalt ions remaining in the spent fermentation medium constitute a serious environmental hazard. However, mutants that do not require xylose and cobalt have been obtained for commercial production of glucose isomerase.

Almost every known glucose isomerase is an intracellular enzyme, and can only be extracted from the cells in relatively low concentrations; so it is expensive to use this enzyme as a soluble and "once only" enzyme.

TABLE 24.7 Glucose Isomerase (GI) Producing Microorganisms

Microorganism	*Patent Assignee*	*Yield (GIU*/Liter)*
Arthrobacter sp.		
NRRL-B-3726	R. J. Reynolds	4,720
NRRL-B-3728		4,440
Streptomyces welmorensis		
ATCC-21175	Standard Brands	4,640/g
Mutant-1		7,540/g
Mutant-2		6,680/g
Mutant-3		6,000/g
Streptomyces olivaceus		
NRRL-3583	Miles	2,560
Mutant NRRL-3916		2,960
Streptomyces olivochromogenes		
Mutant CPC-3	CPC	4,800–11,440
CPC-4		5,700–9,680
CPC-8		3,960–4,440
Actinoplanes missouriensis	Anheuser-Busch	2,500–35,200
Bacillus coagulans	Novo	not known

*1 GIU = The amount of enzyme which converts 1 μmole glucose to fructose/minute.

Immobilized enzyme/cell technology is the key scientific advance that has made the use of glucose isomerase for HFCS production a commercial success. In many industrial cases, whole microbial cells are immobilized by physical means such as entrapment or encapsulation in polymeric materials or by chemical means such as intermolecular cross-linking with glutaraldehyde or covalent binding with diazotized diamino compounds. Commercially, soluble glucose isomerase also is immobilized on DEAE-cellulose. One industrial immobilization process goes as follows: The enzyme-containing cells are homogenized and mixed with glutaraldehyde and inert protein. The resulting gel is shaped into suitable granules, which then are washed and dried.

The immobilized glucose isomerase usually can be used for over 1000 hours at a temperature around 65°C. When the column enzyme activity decreases, the flow rate of the incoming glucose syrup can be adjusted so that the conversion of glucose to fructose is maintained constant.

Proteases

Proteolytic enzymes are by far the most important of the commercially available industrial enzymes. These enzymes, being essential parts of the metabolic system of most living organisms, can be isolated from innumerable sources.

Proteases are classified according to their catalytic mechanism. There are serine proteases, cysteine proteases, aspartic proteases, and metalloproteases. This classification is determined through reactivity toward inhibitors that act on particular amino acid residues in the active site region of the enzyme. The serine proteases are widely distributed among microbes. The enzymes have a reactive serine residue in the active site and are generally inhibited by DFP or PMSF. They are generally active at neutral and alkaline pH. The best-known serine protease is the alkaline protease Subtilisin. The occurrence of cysteine proteases has been reported in only a limited number of fungal species. The proteases are sensitive to sulfhydryl reagents, such as TLCK and iodoactic acid. Aspartic proteases are widely distributed among moulds, including *Aspergillus*, *Penicillium*, and *Rhizopus*, but rarely are found in bacteria and protozoa. They have their maximal activity at low pH, 3 to 4. Many of the aspartic proteases are unstable above neutral pH and are not found in cultures growing at neutral or alkaline pH. Most aspartic proteases are sensitive to epoxy and diazoketone compounds. The pepsin-like aspartic proteases have been used commercialy in processes such as soybean protein hydrolysis; the rennin-like aspartic proteases are capable of clotting milk in a manner similar to animal rennins. The rennin-like proteases from *Mucor* and *Endothia* species have commercial applications in cheese manufacture. The metalloproteases have pH optima between 5 and 9 and are sensitive to metal-chelating agents such as EDTA. Metalloproteases are widespread, and most of the bacterial and fungal metalloproteases contain zinc. Calcium is required to stabilize the protein structure. The best-known metalloprotease is the thermolysin produced by *Bacillus thermoproteolyticus*.

Microbial proteases with widely different properties are produced commercially. *Bacillus* protease, however, represents more than 95 percent of the sale of all proteases. The most important use of *Bacillus* proteases is in detergents. Proteinaceous dirt often precipitates on clothes, and it coagulates during the normal washing process. The addition of proteolytic enzymes to the detergent can easily dissolve proteinaceous stains, which are otherwise difficult to remove.

The alkaline serine protease of *Bacillus licheniformis*, also known as Subtilisin Carlsberg, is the preferred protease in most nonionic and anionic detergents. It attacks many peptide bonds and easily dissolves proteins. It may be used at temperatures up to 65°C, and its pH optimum is close to 9.0, the pH normally used in washing fluids.

A few other proteases also are used in detergents. The serine protease of *Bacillus amyloliquefaciens* has found some applications, presumably because of its substantial content of α-amylase. This may be an advantage for some applications.

Novo Nordisk recently began marketing a bleach-resistant protease (called Durazyme) for use in detergents. In this case, Novo modified the structure of an existing detergent protease by replacing bleach-susceptible amino acid residues with hydrophobic amino acids such as alanine.

Sales of proteases were small and relatively unimportant till about 1965. Then, the use of proteases in detergents created an explosion in the enzyme industry. But in 1971 the backlash came. Allegric symptoms were discovered in some workers handling enzymes in detergent factories. The public, particularly in the United States, was alarmed, and proteases were taken out of most detergents. It was later found that with proper precautions in handling every risk can be eliminated by using proteases in liquid form or by encapsulating the enzymes.

In the tanning industry, alkaline protease from *Bacillus amyloliquefaciens* is used in combination with sulfite for hide treatment and dewooling. A protease from an alkalophilic *Bacillus sp.* is successfully used for dehairing of ox hides in combination with lime, apparently because of the enzyme's stability at pH as high as 12.

In the brewing industry, there is a development toward substitution of malt with unmalted barley and amylase, glucanase, and protease of microbial origin. The neutral protease from *Bacillus amyloliquefaciens* and the therostable neutral protease *Bacillus subtilis var. thermoproteolyticus* have been used by brewers successfully to hydrolyze barley proteins into amino acids and peptides.

Furthermore, *Bacillus* proteases are used in the preparation of protein hydrolyzates. The limited hydrolysis of soybean protein is an especially promising application.

The protease fermentation of the *Bacillus* bacteria takes place under strictly aseptic conditions in conventional equipment for submerged fermentations. The aeration rate is about 1 vvm (volume of air per volume of medium per minute). Vigorous agitation is used to improve air distribution and oxygen transfer. The fermentation temperature is around 37°C, and the time cycle is 2 to 4 days.

The composition of the fermentation medium is important to the yields of protease. Proteins of many different sources are used in commercial media. Carbohydrates are used as an energy source. The C/N ratio is important to the success of the process. Protein should be present in high concentration, and carbohydrate must not be in excess. A convenient way of obtaining this is to conduct fed-batch fermentation, feeding carbohydrate during the run and maintaining the carbohydrate concentration below 1 percent. Continuous fermentation of protease in commercial scale is not yet known.

The recovery and finishing of *Bacillus* protease involves the following steps:

1. Cooling to about 4°C to prevent microbial spoilage.
2. Precipitating undesirable salts using flocculants or filter aids.
3. Removing all particles by centrifugation or filtration.
4. Removing pigments and odors with activated carbon treatment.
5. Removing bacterial contaminants by filtration.
6. Concentrating at low temperature, either by reverse osmosis or by vacuum evaporation.
7. Recovering protease using precipitation by salts (ammonium or sodium sulfate) or solvents (acetone, ethanol or isopropanol).
8. Recovering the precipitate by filtration.
9. Drying at low temperature.
10. Encapsulating granules of enzymes in a nonionic surfactant.

The world production of *Bacillus* protease is estimated to be around 500 tons, with sales close to 100 million dollars.

Proteases also are produced for special applications by many different organisms, such as *Streptomyces griseus*, *Aspergillus niger*, *Aspergillus flavusoryzae*, *Mucor miehei*, *Mucor pusillus*, and *Endothia parasitica*. The applications include protein hydrolysis to amino acids, milk coagulation, cheese manufacturing, and preparation of digestive aids.

Rennin, an aspartyl protease mentioned earlier, is found in the fourth stomach of nursing calves. Because of a decline in veal consumption and an increased demand for cheese, the dairy industry found it difficult to obtain sufficient quantitites of this enzyme from natural source. The increasing market of rennin (around $100 million a year) provided the necessary incentive for commercializing a recombinant DNA-derived chymosin. Prochymosin, the self-processing precursor of chymosin, was expressed in *E. coli* as an insoluble, refractile protein aggregate. The cells were lysed, and the insoluble prochymosin mass was isolated by centrifugation. The prochymosin was solubilized by 8 M urea with high pH treatment, renatured, and activated, and chymosin then was purified by ion exchange chromatography. Pfizer markets the genetically engineered chymosin under the tradename Chymax. Activatable prochymosin also has been expressed in yeast and *Aspergillus* systems.

SCP (SINGLE CELL PROTEIN)

The term "single cell protein" (SCP) was coined at MIT by Professor C. L. Wilson in 1966 to represent the cells of algae, bacteria, yeast, and fungi grown for their protein contents. It should be noted that these microbial cells contain, in addition to proteins carbohydrates, lipids, nucleic acids, vitamins, and minerals.

In the late 1960s and early 1970s, considerable research, development, and commercial interests from all over the world were directed toward SCP production against a background of an increasing deficit in protein supplies, an increasing output, at the time, of cheap petrochemicals, increasing interest in converting waste materials into profitable products, and relatively stable supplies and prices of agricultural products.

Most of this research and industrial effort did not achieve commercialization of SCP, largely because of political, sociological, and psychological factors. However, the SCP era witnessed several important advances in fermentation technology, such as continuous culture, improved bioreactor design, and computer control of fermentation processes.

Most of the SCP produced currently is for animal consumption. Minor quantitites of SCP are being produced for human use. This use has been controversial because of the content of nucleic acid, which can increase blood uric acid levels in a segment of the population. However, many traditional foods contain SCP, such as cheese, yogurt, and sour cream.

The following paragraphs summarize, by substrates, the status of major SCP projects.

Carbon Dioxide for SCP

The photosynthetic *Chlorella* (the green algae) has been grown on carbon dioxide mostly in the Oriental countries. Its market is largely in Japan. The estimated annual consumption is around 20,000 tons.

Methane for SCP

Methane has been of interest as a substrate for SCP production because it is available in large quantity and high purity and, in many cases, is wasted at oil-producing wells. Many bacteria and some fungi utilize methane as a carbon and energy source for growth. A bacterial SCP continuous process was developed to the pilot scale by Shell in England using *Methanomonas capsulatas* and mixed cultures with good productivities and yields. The following factors forced the operation to be discontinued:

- Limited transfer of methane and oxygen from the gas phase to the bacterial cells.
- Explosion hazards.
- High cooling demands because of excessive heat production.
- Presence of inhibitory products.
- High capital requirements because of problems stated above.

n-Paraffins for SCP

The *Candida* yeast grown on *n*-paraffins was widely investigated for production of SCP.

Continuous processes, using air-lift fermentors and operating nonaseptically in most cases, were developed by British Petroleum (using *Candida lipolytica*), Gulf (using *Candida tropicalis*), Kanegafuchi (using *Candida sp.*), Liquichemica Biosintesi (using *Candida maltosa*), and others.

Italproteine, a joint venture of BP and Anic (the petrochemical arm of the Italian state-owned Ente Nazionale Idrocarburi), constructed a 100,000 tons/year plant at Sarroch, Sardinia. Italproteine tradenamed its product "Troprina," made by BP-developed know-how. Liquichimica Biosintesi, owned by Liquichimica (a subsidiary of Italy's Liquigas), built its 100,000 tons/year plant at Saline di Montebello, Calabria. The product was tradenamed "Liquipron," made by the Kanegafuchi process. The operation of these two SCP plants has been blocked by the Italian government's health authorities on questions of levels of hydrocarbon residues.

Plans to construct large-scale yeast SCP plants (60,000–120,000 tons/year each) in Japan by Kanegafuchi, Kyowa Hakko, Asahi, Mitsui Toatsu, and Dainippon were canceled because of public protests over questions on safety of the products and subsequent Japanese government regulatory action banning the use of these products.

However, yeast SCP plants are believed to be in operation in the former Soviet Union using purified *n*-alkanes as substrates. The overall capacity is estimated above 1,500,000 tons/year. The increasing production of SCP in the former Soviet Union as feed ingredients reduces the need to import grains.

Methanol for SCP

The events in Italy and Japan, as well as the dramatic increases in petroleum prices in the mid-1970s, stimulated much increased effort in selecting substrates other than *n*-paraffins for SCP production. Methanol has received perhaps the greatest attention. The use of methanol as a substrate for SCP production has a number of advantages, principally the following:

- High solubility in water.
- Low explosive hazards of methanol oxygen mixtures.
- Moderate heat liberation (3.3–5.7 kcal/g of cells vs. 4.4–8.0 kcal/g of cells grown on *n*-paraffins).
- Freedom from traces of aromatic hydrocarbons.
- Ease of removal from the microbial cell product.

Organizations in many parts of the world have worked on methanol processes using both yeasts and bacteria, including Shell and ICI in England, Societa' Italiane Resine in Italy, Hoechst AG and Gelsenberg in Germany, Norsk Hydro in Norway, AB Marabou in Sweden, Yissum Research Development in Israel, Phillips Petroleum and Tenneco in the United States, and Mitsubishi in Japan. In addition, much work has been done in academic institutions. Among the industrial organizations, ICI and Hoechst are the ones that remain most active. Both use bacterial processes that have the advantages over yeast processes of higher growth rates, productivities, and yields.

Hoechst operates a 1000 tons/year pilot plant near Frankfurt. *Methylomonas clara* is grown on a mixture of air and methanol containing ammonia, water, and other essential mineral nutrients. The dried product, tradenamed Probion, contains some 70 percent protein, 10 percent nucleic acid, 8 percent fats, 7 percent mineral "ash," and 5 percent water. A purified version, with essentially all nucleic acids and fats removed, has a protein content of more than 90 percent, with minerals and water making up the balance. Hoechst intends to market this purified product, called Probion-S, for human consumption.

ICI has developed the most advanced process for SCP production from methanol. *Methylophilus methylotrophus* is the organism used. A novel type of air-lift bioreactor, called the pressure cycle fermentor, has been developed for use in the process. It is designed to maintain a high oxygen transfer rate and a homogeneous liquid phase without excessive shear, to remove the heat liberated during the growth phase at high productivities, and to

avoid contamination problems encountered with conventional stirred fermentors because of leaks through drive shafts and mechanical seals. Air and fresh medium are introduced at the base of a tall vertical column—the "riser." The riser is connected horizontally to another smaller diameter column, the "downcomer," at both the top and the base. A driving force for the movement of broth at high velocities is provided by the difference in air holdup between the riser and the downcomer. The high hydrostatic head and the level of turbulence in the riser provide excellent oxygen transfer because of the increased pressure, the small bubble size, and the mixing; up to 50 percent of the incoming oxygen is transferred. Spent air and CO_2 (stripped by the nitrogen content of the air) are disengaged in the upper horizontal section. The broth is cooled in a simple heat exchanger set in the downcomer. ICI also has developed a proprietary agglomeration process for the initial separation of bacterial cells from the growth medium, which permits the final centrifugation of a much higher solid slurry than is possible otherwise.

A 70,000 tons/year plant came on stream in early 1980. A schematic diagram of the ICI protein process is shown in Fig. 24.30. The final product, tradenamed Pruteen, contains more than 70 prcent protein and is sold as animal feed. The selling price is about $500/ton. The ICI pressure cycle fermentor at Billingham, England, measuring 200 feet high and 25 feet in diameter, is claimed to be the largest such vessel in the world.

The naturally occurring *Methylophilus methylotrophus* used in the ICI SCP process is relatively inefficient in assimilating nitrogen. The inefficiency is caused by the absence of the gene that codes for glutamate dehydrogenase, the enzyme essential for incorporating hydrogen and ammonia into glutamic acid, the major pathway for the formation of α-amino groups directly from ammonia. Instead, *Methylophilus methylotrophus* possess the gene for producing glutamate synthase, which leads to a less efficient nitrogen assimilation pathway. It is reported that ICI has successfully overcome the problem by first inactivating the glutamate synthase enzyme using classical mutation techniques and then inserting into *Methylophilus methylotrophus* the glutamate dehydrogenase gene from *Escherichia coli.*

Ethanol for SCP

Ethanol can be used by certain bacteria and yeasts as a substrate for SCP production. It is completely water-soluble, available in large quantity and high purity, and easy to store

Fig. 24.30. Flowsheet of the ICI SCP process. (Process Biochem., *12, (1), 30, 1977.*) (Reprinted with permission of the publisher.)

and transport. In addition, it has the advantage over methanol of low toxicity and good acceptability to regulatory agencies as a raw material for producing a food-grade end product. Organisms of interest that utilize ethanol in SCP production include:

1. *Candida utilis* (used by Pure Culture Products, formerly Amoco Foods).
2. *Acinetobacter calcoaceticus* (used by Exxon and Nestle).
3. *Candida acidothermophilum* and *Candida ethanothermophilum* (used by Mitsubishi Petrochemical).

The yeast-based SCP processes using ethanol as a substrate have been developed to a greater extent than the bacterial processes.

Pure Culture Products in the United States has taken the approach of using torula yeast, which is well known and has long been established as a food ingredient. As such, they have been able to obtain clearance for their ethanol-grown product as a food ingredient, and they are already marketing material from their plant in Hutchinson, Minnesota, which is ultimately designed to produce up to 15 million lb of product per year. The material, tradenamed Torutein, is being sold in the United States as a nutritional supplement and flavor enhancer for such processed foods as meat patties, pasta, baked goods, frozen pizzas, and sauces.

In the Pure Culture Product's continuous process, the ethanol concentration is maintained at approximately 200 ppm. The nitrogen source, either aqueous or anhydrous ammonia, is supplied continuously to keep the pH in the desired range. Other macro and micro elements also are supplied continuosuly. All liquid streams except ammonia are sterilized at 149°C. Ammonia is sterilized by filtration, and air is sterilized by compression and filtration. Typical operating conditions are 30°C, pH 4.6, and aeration and agitation to give oxygen absorption rates in the range of 100 to 140 millimoles of O_2 per liter per hour. Cell concentrations of 6 to 7 g/liter on a dry weight basis are obtained at a dilution rate of about 0.3 hr^{-1}.

Carbohydrates for SCP

The raw materials used for SCP production in this category are carbohydrate-containing wastes or by-products of various manufacturing processes. These carbohydrate-containing materials include molasses, sulfite liquor waste, whey, and cellulosic wastes.

Yeasts grown on molasses are produced for both food and feed uses by Standard Brands, Universal Food, and Yeast Products in the United States, and by many other processors abroad.

Sulfite liquor waste is used for feed yeast production by Boise Cascade, Rayonier (a subsidiary of ITT), and Wausau Paper Mills in the United States, by Environcon in Canada, by Attishola in Switzerland, and by G. A. Serlacins in Finland. Yeast plants utilizing sulfite liquor or wood hydrolyzate are also in operation in Austria, Rumania, Poland, Germany, Japan, the People's Republic of China, Saudi Arabia, and the former Soviet Union.

One well-publicized SCP process using spent sulfite liquor as a substrate is the Pekilo process developed by Tampella in Finland. The Pekilo process is said to be the first commercial continuously operating process in which filamentous fungi are cultivated for SCP production. A plant based on the Pekilo process was built jointly by United Paper Mills and Tampella at Jamsankoski in Finland with a capacity of 10,000 tons/year using a microfungus *Paecilomyces varioti*. Its product, the Pekilo protein, is sold for animal feed.

Whey, a waste of cheese manufacturing, is fermented with *Saccharomyces fragilis* by Amber Laboratories, Stauffer Chemical, and Kraft in the United States for both food and feed uses. For the production of food yeast, the yeast cream obtained by centrifugation is spray-dried. For feed yeast, the whole fermentor broth is concentrated by evaporation and then spray-dried.

Cellulose and cellulosic wastes have received a great deal of attention in recent years as potential substrates for SCP production. Delignification, decrystallization, hydrolysis, and microbial growth are the important steps

involved in SCP production from cellulosic materials. At the present time, no commercial production of SCP from cellulose is known. However, the utilization of cellulose, the most abundant renewable organic resource on earth, is of great importance from both technical and commercial points of view, not only for SCP production but also for chemicals and fuels.

ANTIBIOTICS

The science of microbiology was established by the French chemist Louis Pasteur. Among his many contributions, the discovery of the existence of microscopic forms of life should be especially pointed out: yeasts were established as the cause of the alcohol fermentation of wine. He later established pathogenic bacteria as the cause of many infectious diseases and even recognized viruses as causative agents of other infections.

Shortly before World War II, another significant role of microorganisms began to emerge. Alexander Fleming, an English bacteriologist, is credited with first suggesting that the product of one microorganism might be used to inhibit the growth of another. Fleming observed that a chance contaminant (a *Penicillium* mold) clearly prevented the growth of a pathogenic *Staphylococcus* he was culturing in a Petri dish. He succeeded in establishing some of the simple properties of the mold product penicillin and published his results in 1929. Nearly a decade later, another group of English biochemists undertook to further examine the phenomenon in the course of a broad research program for better chemotherapeutic agents. Florey, Chain, Heatley, and Abraham, of Oxford University, succeeded in isolating a small quantity of penicillin concentrate and by 1941 unequivocally demonstrated its potential usefulness in *Staphylococcus* septicemia. British government officials recognized its possible usefulness in military medicine, but could not further develop the discovery because of the serious war conditions then prevailing. The research was brought to the attention of the government and industry in the United States, with the result that an international government–industry research program was established to produce this remarkable chemotherapeutic agent. The success of this program established a new class of powerful therapeutic agents, the antibiotics, which have revolutionized medical practice. Thus the product of one microorganism is used to combat an infection caused by another. To supply the huge amounts of antibiotics needed in modern medicine, the fermentation industry has undergone a virtual revolution.

Since the early 1940s, an intensive search for new and useful antibiotics has been in progress throughout the world. In a period of a little more than 35 years, over 4300 antibiotics have been discovered from microbial origins. More than 30,000 semisynthetic antibiotics have been prepared. Of the thousands of antibiotics known, over 110 are produced on a commercial scale, with approximately one half of them prepared by a combination of microbial synthesis and chemical modification.

Penicillins, cephalosporins, and tetracyclines are the most important antibiotic groups in terms of production tonages and dollar sales. The 1990 U.S. sales of cephalosporins were close to $2 billion, penicillins (including penicillin/beta-lactamase inhibitor combinations) $900 million, and tetracyclines (as human therapeuics) $200 million. The other important human antibiotics include macrolides (mainly erythromycin), aminoglycosides (such as neomycin and gentamycin), streptomycin, vancomycin, clindamycin, lincomycin, bacitracin, and monobactams.

Penicillins

The original mold observed and preserved by Fleming was a strain of *Penicillium notatum*, a common laboratory contaminant. Later cultures of *Penicillium chrysogenum* were found to be better producers of penicillin, and the present industrial strains have been derived from this species. The original strains produced the antibiotic only by surface fermentation methods and in very low yields,

a few ppm. Gradually, improved media and the eventual discovery of strains productive under submerged aerobic fermentation conditions led to dramatic yield increases, which made commercial production a reality. Subsequent improvements, principally in culture selection and mutation to productive strains, further improved yields, until today broths often contain 20 to 30 g/liter of penicillin. Dramatic price reduction has come with improved production. For instance, a million-unit vial of penicillin (1667 units = 1 mg of potassium penicillin G) had a wholesale price of $200 in 1943. In 1952, the same vial had a wholesale value of only $1.30. Today, a million units of potassium penicillin G is selling for as little as $0.50 or approximately $20/kg of free acid.

The original *P. chrysogenum* strains produced large amounts of unwanted yellow pigments that were difficult to remove from the recovered penicillin. Today, nonpigmented mutants, a strain known in the industry as Wisconsin 49–133 (or progeny therefrom), are universally employed. The desired culture is propagaged from a laboratory stock in small flasks and transferred to plant inoculum tanks. After 24 hours these are used to inoculate larger fermentors, which contain a typical production medium such as that shown below.

Components	*Grams/Liter*
Corn steep liquor	30
Lactose	30
Glucose	5.0
$NaNO_3$	3.0
$MgSO_4$	0.25
$ZnSO_4$	0.044
Phenyl acetamide (precursor)	0.05
$CaCO_3$	3.0

The medium is usually sterilized batchwise, cooled to 24°C, and inoculated. The time of fermentation may vary from 60 to 200 hours. Sterile air is blown through the tank, usually at a rate of one volume per minute.

When penicillin concentration reaches its peak potency, as determined by microbiological or chemical assays, the broth is clarified by means of rotary vacuum filters. The penicillin, being acidic, is extracted from the aqueous phase into a solvent, such as methyl isobutyl ketone or amyl acetate, at a pH of 2.5 by means of a continuous countercurrent extractor, such as a Podbielniak. The penicillin extract then is re-extracted with an aqueous alkaline solution or a buffer at a pH of 6.5 to 7.0. A 90 percent recovery is made at this step. The aqueous solution is chilled, acidified, and extracted again with a solvent, such as ether or chloroform. The solvent extract then is reextracted into water at a pH of 6.5 to 7.0 by titration with a solution of base. The base used depends on which salt of penicillin is desired. The popular forms are sodium or potassium salts. A typical flow sheet for antibiotic recovery is shown in Fig. 24.15.

Table 24.8 gives the structural formulae of the "natural penicillins," comprising several closely related structures with aliphatic and aromatic substitutions to the common nucleus. The early impure product contained mixtures of these types. For several reasons penicillin G became the preferred type and the crystalline product of commerce. Phenylacetic acid or its derivatives are used as precursors in the fermentation medium to enhance penicillin biosynthesis and suppress the production of the less desirable types.

The fact that proper selection of precursors could lead to new variations in the penicillin side chain offered the first source of synthetic penicillins. Penicillin V, derived from a phenoxyacetic acid precursor, attracted clinical use because of its greater acid tolerance, which made it more useful in oral administration.

The widespread use of penicillin eventually led to a clinical problem of penicillin-resistant staphylococci and streptococci. Resistance for the most part involved the penicillin-destroying enzyme, penicillinase, which attacked the beta-lactam structure of the 6-amino-penicillanic acid nucleus (6-APA).

In 1959, Batchelor and coworkers in the Beecham Research Laboratories in England discovered that the penicillin nucleus, 6-APA, accumulated during fermentation when side chain precursors were omitted. This 6-APA

TABLE 24.8 Structural Formulae of Natural Penicillins

Formula:

Type of Penicillin	*Side Chain R Substitutions*
(G) Benzyl	
(X) *p*-Hydroxybenzyl	
(F) 2-Pentenyl	$CH_3—CH_2—CH=CH—CH_2—$
(Dihydro F) *n*-Pentyl	$CH_3—CH_2—CH_2—CH_2—CH_2—$
(K) *n*-Heptyl	$CH_3—CH_2—CH_2—CH_2—CH_2—CH_2—CH_2—$
(V) Phenoxy	

could be used for the chemical synthesis of entirely new types of penicillin by coupling with new side chains. Shortly thereafter, several sources of penicillin amidase were found that would cleave the phenylacetyl side chain from penicillin G, thus producing a more economical source of 6-APA. A vast number of "synthetic penicillins" have been generated, and a few have achieved clinical importance. Several objectives were sought:

1. To broaden the inherent utility of penicillin to include gram-negative pathogens not inhibited by the natural penicillins.
2. To improve its stability and absorption.
3. To increase its resistance to penicillinase-producing pathogens.
4. To decrease allergenicity.
5. To improve other factors pertient to clinical use.

The broad objectives have been achieved with varying degrees of success.

Table 24.9 shows the structures of some of the semisynthetic penicillin that have become important chemotherapeutics. Semisynthetic penicillins, on the average, sell for about $150/kg.

The natural penicillins, primarily G and V, have a relatively narrow spectrum. They act primarily on gram-negative organisms. Semisynthetic penicillins such as ampicillin, carbenicillin have a broader spectrum. Some, such as methicillin, oraficillin, and oxacillin, are resistant to penicillinase. In 1984, Beechem introduced Augmentin, which was the first combination formulation of a penicillin (amoxillin) and a penicillinase inhibitor (clavulanic acid). In 1985, Beechem introduced Timentin, a ticarcillin/clavulanic acid combination, and Pfizer in 1987 introduced Unasyn, an ampicillin/sulbactam sodium combination.

Cephalosporins

In 1948 Professor Guiseppe Brotzu isolated a *Cephalosporium* culture from seawater near the sewer discharge of Cagliarri, Sardinia. This culture produced a broth inhibiting both gram-positive ad gram-negative bacteria. The team of Florey, Abraham, and Newton at Oxford University isolated a compound identified as cephalosporin N; and during the same period a group in the Michigan Department of Health isolated synnematin B from another strain of *Cephalosporium*. Synnematin B and cephalosporin N proved to be identical. Structure studies eventually proved these antibiotics to be a new type of penicillin, α-aminoadipyl-6-APA (also called

TABLE 24.9 Structural Formulae of Some Semisynthetic Penicillins

Carbenicillin		Nafcillin	
Penicillin V		Quinacillin	
Phenethicillin		Oxacillin	
Propicillin		Cloxacillin	
Phenbenicillin		Dicloxacillin	

Methicillin		Flucoxacillin	
Ampicillin		Azidocillin	
Epicillin		Cyclacillin	
Amoxicillin		Metapilicillin	
Hetacillin		Mecillinam	

penicillin N), a naturally produced penicillin with gram-negative activity.

In the course of studies on the Brotzu strain of *Cephalosporium*, Abraham and Newton detected small quantities of a second antibiotic, cephalosporin C. Painstaking work proved it to be chemically similar to penicillin N, but not a penicillin. It had pronounced gram-negative activity, was more stable to acid, and was not destroyed by penicillinase. It possessed the same α-aminoadipyl side chain as a new penicillin, but the nucleus was 7-aminocephalosporanic acid (7-ACA). 7-ACA contains a six-membered 1,3-dihydrothiazine ring instead of the five-membered thiazole ring in 6-APA. The structures of 6-APA and 7-ACA are shown in Fig. 24.31.

Microbiological processes for production of cephalosporin C resemble in many respects those used for penicillin production. Special strains of *Cephalosporium* have been selected that produce more cephalosporin C and less cephalosporin N than the parent culture. The growth of these strains in certain special fermentation media has resulted in higher antibiotic titers. Even with these improvements in processing, the antibiotic yields, averaging 10 to 20 g/liter, are much lower than those reported for the penicillins. Cephalosporins sell for around $250/kg.

As in the penicillin studies, the possibility of further improving the chemotherapeutic properties of cephalosporin C was apparent if the 7-ACA nucleus could be obtained. Enzymatic cleavage of the side chain failed, as did the use of precursors to generate new side chains; however, successful chemical methods have been found. Several semisynthetic cephalosporins have been produced and are used clinically.

Fig. 24.31. Structures of 6-APA and 7-ACA.

The leading agents in this category include cefaclor (first-generation cephalosporin, developed by Lilly), cefoxitin (second-generation, by Merck), cefuroxime axetil (oral form, second-generation, by Glaxo), and ceftriaxone (third-generaton, by Roche).

The enzymatic removal of the 6-acyl side chain of penicillins is widely used in the commercial production of 6-APA. Bacterial acylases also have been found that cleave the 7-acyl side chain of chephalosporin C to form 7-ACA. The discovery and the development of penicillin and cephalosporin acylases greatly accelerated the commercialization of many semisynthetic penicillins and cephalosporins. On the other hand, basic studies on biosynthesis of beta-lactam antibiotics have progressed rapidly with the recent applications of molecular biology techniques to many of the beta-lactam–producing organisms. Most of the biosynthetic reactions involved in the formation of penicillins and cephalosporins have been demonstrated in cell-free systems, and a few of the enzymes have been highly purified and cloned. Knowledge of these enzymes, their promoters, and efficient expression will have a positive influence on the industrial production of penicillins and cephalosporins.

Tetracyclines

In 1948, a broad-spectrum antibiotic, chlortetracycline (Aureomycin), was announced from the Lederle Laboratories, Division of American Cyanamid Company. This antibiotic is produced by *Streptomyces aureofaciens* when grown under submerged aerobic conditions on media composed of sugar, cornsteep liquor, and mineral salts. The crystalline compound has a golden yellow color, which suggested the trade name.

The following year a second related antibiotic, oxytetracycline (Terramycin), a product of *Streptomyces rimousus*, was announced by Pfizer Inc. It also is a yellow substance, chemically and biologically similar to

chlortetracycline. Independent research by both companies eventually led to the disclosure of the structure of these two important chemotherapeutic agents; this has been regarded as one of the brilliant achievements of modern organic chemistry.

Both compounds may be regarded as derivatives of a nucleus known as tetracycline. Their structures along with those of other clinically important tetracyclines are shown in Table 24.10.

Tetracycline also can be produced by *Streptomyces aureofaciens* fermentations under special conditions, that is, chloride starvation or special strains of the organism that fail to halogenate efficiently. Tetracycline possesses many chemotherapeutic properties of chlortetracycline and oxytetracycline. It is an important broad-spectrum antibiotic.

Mutations of tetracycline-producing organisms have led to other tetracycline analogues, of which 6-demethyl-7-chlortetracycline (Declomycin) has clinical use. Chemical modifications of oxytetracycline have generated two other useful members of the family, known as Vibramycin and Rondomycin.

Tetracyclines are active *in vivo* against numerous gram-positive and gram-negative organisms, and some of the pathogenic rickettsiae and large viruses. The systemic administration of tetracyclines may be carried out by utilizing either oral or intravenous dosage forms. For veterinary use, tetracyclines are given by intravenous injection in bovines and equines, except that oral routes are used for young nonruminating calves.

In connection with work on animal protein factors, it was found that chlortetracycline fermentation mash containing some vitamin B_{12} gave growth responses well above those obtained with supraoptimal levels of vitamin B_{12} alone. At the present time, both chlortetracycline and oxytetracycline are used extensively for growth stimulation and improvement of feed efficiency in poultry and hogs, and for the reduction of losses from certain disease conditions. Purified antibiotics as well as dried fermentation residues with the mycelium of *Streptomyces aureofaciens* are used for these purposes. Tetracyclines, used as animal feed supplements, are selling for about $60/kg.

TABLE 24.10 Structure of Clinically Important Tetracyclines

	R_1	R_2	R_3	R_4
Tetracycline	H	OH	CH_3	H
7-Chlortetracycline (Aureomycin®)	H	OH	CH_3	Cl
5-Oxytetracycline (Terramycin®)	OH	OH	CH_3	H
6-Demethyl-7-chlortetracycline (Declomycin®)	H	OH	H	Cl
6-Deoxy-5-oxytetracycline (Vibramycin®)	OH	H	CH_3	H
6-Methylene-6-deoxyl-6-demethyl-5-oxytetracycline (Rondomycin®)	OH	H	CH_2	H

Other Antibiotics

Table 24.11 lists some antibiotics produced on a commercial scale. The importance of new antibiotics for use in clinical medicine is stressed in most discussions of antibiotics, but their importance in other areas should not be forgotten. Antibiotics such as streptomycin and tetracycline are being used against bacterial plant pathogens, and cycloheximide, blasticidin S, nystatin, and griseofulvin are being used against fungi. Antibiotics also are being used in livestock production, where they improve marketable weight and increase food utilization. Those that stimulate animal growth are bacitracin for poultry and swine; bambermycin for swine and calves; virginamycin for poultry and swine; avoparcin for poultry and swine; chlortetracyline for poultry, swine, calbes, cattle, and sheep; erythromycin for chickens; nystin, oleandomycin, and procaine penicillin for poultry and swine; streptomycin for poultry; oxytetracycline for poultry, swine, cattle, and calves; tylosin for poultry and swine; lasolocid for poultry; and monensin for poultry and cattle. Combinations of antibiotics and their mixtures with sulfa drugs also are being used.

Further research on new agents to treat both human and animal diseases is certainly in order, particularly to treat those diseases that are not successfully controlled at present.

Native Bioactive Products

In the five decades following the first clinical application of penicillin, the investigation of microorganisms as sources of therapeutically useful compounds has focused on those with antibiotic activity against bacterial and fungal infections. Indeed, antibiotics discovered in this period have been the basis of a major advance in the practice of medicine. In the last ten years, however, it has become evident that the contribution of microorganisms does not have to be limited to the realm of infectious diseases. Metabolites of microorganisms have been found to have many other interesting therapeutic applications. Particularly as our understanding of disease processes at the biochemical and genetic levels has advanced, enzymes or receptors have served as specific targets for therapeutic intervention. Recombinant DNA technology, in particular, has helped to produce reagents as tools for the development of practical and high-throughput screening methodologies based on the mechanism of action. Several of these natural bioactive producers have been successfully developed as clinically useful therapeutics.

Cyclosporin A, a fungal decapeptide, is the principal drug used to suppress immune rejection in organ transplant patients. Cyclosporin A binds to a cytosolic protein, cyclophilin, inhibits its peptidyl-prolyl isomerase activity, and suppresses T cell activation. Cell-based screens for fermentation products with similar immunomodulating activity have resulted in compounds with biological activities similar to that of cyclosporin A. Fujisawa's FK506, the most interesting among them, also binds to a cytosolic protein, FKBP (FK506 binding protein). It too inhibits peptidyl-prolyl isomerase activity, and suppresses T cell activation. FK506 has been proved as an effective immunosuppressant and has fewer side effects than cyclosporin A.

Mevinolin and compatin, independently discovered by Merck and Sankyo, both control cholesterol synthesis in humans. The Merck group screened directly for inhibition of HMG-CoA reductase, a key enzyme involved in cholesterol synthesis, while the Sankyo group sought agents that would interfere the incorporation of radiolabelled acetate into cholesterol in a cell-free enzyme system. The search for microbial products that interfere with cholesterol metabolism continues, given the importance of serum cholesterol in coronary heart disease. Compatin and mevinolin, being HMG-CoA reductase inhibitors, also prevent maturation (a post-translational modification near the carboxy terminal) of *ras* proteins. They thus may be of value in the treatment of *ras*-dependent tumors.

Monobactam, a new class of natural beta-lactam antibiotics, was found from a beta-lactamase screen. Aztreonam, a chemically synthesized monobactam, is in clinical use.

TABLE 24.11 Some Antibiotics Produced on a Commercial Scale*

		Antibiotic spectrum[b]							
Antibiotic	*Microbial source*	*G+*	*G−*	*My*	*AF*	*AT*	*Other*	*Chemical type*	*Therapeutic or other use*
Amphomycin	*Streptomyces canus*	+						Peptide	Topical
Amphotericin B	*Streptomyces nodosus*				+			Polyene	Oral or parenteral
Avoparcin[a]	*Streptomyces candidus*	+	+					Glycopeptide	Animal growth promotant
Azalomycin F[a]	*Streptomyces hygroscopicus*	+	+		+				Topical (AF)
Bacitracin	*Bacillus subtilis*	+						Peptide	Topical; also animal growth promotant
Bambermycins	*Streptomyces bambergenesis*	+						Phosphoglycolipid	Animal growth promotant
Bicyclomycin[a]	*Streptomyces sapporonensis*	+	+						Topical
Blasticidin S[a]	*Streptomyces griseochromogenes*				+			Nucleoside	Agricultural (AF)
Bleomycins	*Streptomyces verticillus*	+	+			+		Peptide	Parenteral (AT)
Cactinomycin[a]	*Streptomyces chrysomallus*	+				+		Peptide	Parenteral (AT)
Candicidin B	*Streptomyces griseus*				+			Polyene	Topical
Candidin[a]	*Streptomyces viridoflavus*				+			Polyene	Topical
Capreomycin	*Streptomyces capreolus*			+				Peptide	Parenteral
Cephalosporins	Cephalosporin C is produced by *Cephalosporium acremonium* and converted to 7-ACA, which is used for prep of semisynthetic cephalosporins	+	+					Peptide	Oral and parenteral
Chloramphenicol	*Streptomyces venezuelae*; commercial manufacture is by chemical synthesis	+	+				Rickettsia		Oral or parenteral
Chromomycin A_3[a]	*Streptomyces griseus*	+				+			Parenteral (AT)
Colistin	*Bacillus colistinus*		+					Peptide	Parenteral
Cycloheximide	*Streptomyces griseus*				+				Agricultural (AF)
Cycloserine	*Streptomyces orchidaceus*			+				Amino Acid	Parenteral (TB)
Dactinomycin	*Streptomyces antibioticus*					+		Peptide	Parenteral (AT)
Daunorubicin	*Streptomyces peucetius*					+			Parenteral (AT)
Doxorubicin	*Streptomyces peucetius*					+			Parenteral (AT)
Enduracidin[a]	*Streptomyces fungicidus*	+		+				Peptide	Animal growth promotant
Erythromycin	*Streptomyces erythreus*	+						Macrolide	Oral and parenteral; animal growth promotant
Fortimicins[a]	*Micromonospora olivoasterospora*	+	+					Aminoglycoside	Parenteral
Fungimycin[a]	*Streptomyces coelicolor* var *aminophilus*				+			Polyene	Topical
Fusidic Acid[a]	*Fusidium coelcineum*	+						Steroid	Parenteral
Gentamicins	*Micromonospora purpurea*	+	+					Aminoglycoside	Parenteral

Gramicidin A	*Bacillus brevis*	+						Peptide	Topical
Gramicidin J[a](S)	*Bacillus brevis*	+						Peptide	Topical
Griseofulvin	*Penicillium griseofulvum*				+			Spirolactone	Oral
Hygromycin B	*Streptomyces hygroscopicus*	+	+				Helminths	Aminoglycoside	Animal feed suppl.
Josamycin[a]	*Streptomyces narbonesis*	+						Macrolide	Oral and parenteral
Kanamycins	*Streptomyces kanamycetcus*	+	+	+				Aminoglycosides	Parenteral
Kasugamycin[a]	*Streptomyces kasugaensis*	+	+					Aminoglycoside	Agricultural antibacterial
Kitasatamycin[a]	*Streptomyces kitasatoensis*	+						Macrolide	Oral and parenteral
Lasalocid	*Streptomyces hazelensis*	+					Coddidia	Polyether	Agricultural use as coccidiostat and growth promotant
Lincomycin	*Streptomyces lincolnesis*	+							Oral and parenteral
Lividomycin[a]		+	+						
Macarbomycins[a]	*Streptomyces phaechromogenes*	+						Phosphoglycolipid	Animal growth promotant
Mepartricin[a]					+			Polyene	Topical
Midecamycin[a]		+						Macrolide	Oral and topical
Mikamycins[a]	*Streptomyces mitakaensis*	+						Peptide	Animal growth promotant
Mithramycin	*Streptomyces* species	+		+		+			Parenteral (AT)
Mitomycin C	*Streptomyces caespitosus*	+	+		+				Parenteral
Mocimycin[a]		+							Animal growth promotant
Monensin	*Streptomyces cinnamonensis*	+					Coccidia	Polyether	Animal growth promotant
Myxin	*Chromobacterium iodinium* plus chemical modification				+			Phenazine	Topical in veterinary use
Neocarzinostatin	*Streptomyces carzinostaticus*	+				+		Peptide	Parenteral (AT)
Neomycins	*Streptomyces fradiae*	+	+					Aminoglycoside	Oral and topical
Nosiheptide[a]	*Streptomyces actinosus*	+						Peptide	Animal growth promotant
Nisin	*Streptomyces cremoris*	+						Peptide	Food preservative
Novobiocin	*Streptomyces niveus*	+							Oral and topical
Nystatin	*Streptomyces noursel*				+			Polyene	Oral and topical
Oleandomycin	*Streptomyces antibioticus*	+						Macrolide	Oral and parenteral
Paromomycin	*Streptomyces rimosus*	+	+				Protozoa	Aminoglycoside	Oral
Penicillin G[a]	*Penicillium chrysogenum*	+						Peptide	Oral and parenteral; also as animal growth promotant
Penicillin V[a]	*Penicillium chrysogenum*	+						Peptide	Oral
Pimaricin	*Streptomyces natalensis*					+		Polyene	Topical; also used for food preservation
Polymyxin B	*Bacillus polymyxa*				+			Peptide	Parenteral
Polyoxins[a]	*Streptomyces cacaoi* var. *asoensis*								Agriculture (AF)
Pristinamycins[a]	*Streptomyces pristinaspiralis*	+						Peptide	Parenteral
Quebemycin[a]	*Streptomyces viridans*	+						Phosphoglycolipid	Animal growth promotant
Ribostamycin[a]	*Streptomyces ribosidificus*	+	+					Aminoglycoside	Parenteral
Rifamycin SV[a]	*Nocardia mediterranei*			+				Anasamycin	Parenteral

TABLE 24.11 Some Antibiotics Produced on a Commercial Scale* (*Continued*)

		Antibiotic spectrum[b]							
Antibiotic	*Microbial source*	*G+*	*G−*	*My*	*AF*	*AT*	*Other*	*Chemical type*	*Therapeutic or other use*
Ristocetin	*Nocardia lurida*	+						Glycopeptide	Parenteral
Sagamycin[a]	*Micromonospora sagamiensis*	+	+					Aminoglycoside	Parenteral
Salinomycin[a]	*Streptomyces albus*	+					Coccidia	Polyether	Veterinary use
Siccanin[a]	*Helminthosporium siccans*				+				Veterinary use
Siomycin	*Streptomyces sioyaensis*	+						Peptide	Animal growth promotant
Sisomicin	*Micromonospora inyoensis*	+	+					Aminoglycoside	Parenteral
Spectinomycin	*Streptomyces spectrabilis*	+						Aminocyclitol	Patenteral
Spiramycin[a]	*Streptomyces ambofaciens*	+						Macrolide	Parenteral and oral
Streptomycin	*Streptomyces griseus*	+	+	+				Aminoglycoside	Parenteral; use in agriculture to control bacteria
Dihydrostreptomycin	*Streptomyces humidus* (Also chemical reduction of streptomycin)	+	+	+				Aminoglycoside	Parenteral
Tetracyclines									
Chlortetracycline	*Streptomyces auerofaciens*	+	+				Rickettsia	Tetracycline	Parenteral and oral; animal growth promotant
6-Demethyl-7-chlortetracycline	*Streptomyces aureofaciens*	+	+				Rickettsia	Tetracycline	Parenteral and oral
5-Hydroxytetracycline	*Streptomyces rimosus*	+	+				Rickettsia	Tetracycline	Parenteral and oral animal growth promotant
Tetranactin[a]	*Streptomyces flaveolus*	+					Insects	Macrotetralide	Insecticide
Thiopeptin[a]	*Streptomyces tateyamensis*	+						Peptide	Animal growth promotant
Thiostrepton	*Streptomyces azureus*	+						Peptide	Animal growth promotant
Tobramycin[a]	*Streptomyces tenebrarius*	+	+					Aminoglycoside	Parenteral
Trichomycin[a]	*Streptomyces hachijoensis*				+		Trichomonas	Polyene	Topical
Tylosin	*Streptomyces fradiae*	+					PPLO	Macrolide	Veterinary; animal growth promotant
Tyrothricin	*Bacillus brevis*	+	+					Peptide	Topical
Tyrocidine	*Bacillus brevis*	+	+					Peptide	Topical
Validamycin[a]	*Streptomyces hygroscopicus* var. *limonesis*	+	+					Aminoglycoside	Parenteral
Vancomycin	*Streptomyces orientalic*	+						Glycopeptide	Parenteral
Variotin[a]	*Paecilomyces varioti*				+				Topical
Viomycin	*Streptomyces floridae*	+	+	+				Peptide	Parenteral
Virginiamycin	*Streptomyces virginiae*	+						Peptide	Animal growth promotant

[a]Not distributed in the United States (1978).
[b]G+, gram-positive bacteria; G−, gram-negative bacteria; My, mycobacteria; AF, antifungal; AT, antitumor.
**Microbial Technology*, 2nd ed., Vol. 1, p. 244, Academic Press, New York, 1979.

Thienamycin also was discovered in an assay based on beta-lactamase susceptibility. Mutastein, a fungal glycoprotein that inhibits the synthesis of insoluble glucan by the bacteria causing the formation of dental plaque, has been used commercially as a prophylactic agent for tooth decay.

The search for enzyme- or receptor-targeted microbial products does not limit itself to medical use. Several commercially important "antibiotics" of agricultural applications have been found in recent years. Nikkomycin, a chitin synthase inhibitor, is an antifungal agent. Avermectin, acting upon the synaptic gamma-aminobutyric acid (GABA) receptor system, is an effective antiparasitic and anthelmintic agent.

In the past decade, as therapeutic screening became increasingly mechanism-based, production of target enzymes or receptors using recombinant DNA techniques expanded drastically the number of targets that can feasibly be screened. The following table highlights some of the targets and their applications.

Targets	*Applications*
HIV reverse transcriptase	AIDS therapy
HIV protease	AIDS therapy
IL-1 inhibition	Antiarthritic
IL-1 receptor antagonist	Antiarthritic
MHC antagonist	Antiarthritic
Collagenase inhibitor	Antiarthritic
	Anticonvulsant
Monoamine oxidase A inhibitor	Antidepressant
5-HT receptor agonist	Antidepressant
5-HT reuptake inhibitor	Antidepressant
	Antiobesity
Lipase inhibitor	Antiobesity
H2-receptor antagonist	Antiulcerant
Enkephalinase inhibitor	Antiulcerant
	Anagestics
Opioid receptor kappa agonist	Analgesics
Alpha glucosidase inhibitor	Antidiabetes
H1-receptor antagonist	Antihistamine
5-Lipoxygenase inhibitor	Antiasthmatics
Phosphodiesterase inhibitor	Antiasthmatics
PAF antagonist	Antiasthmatics
IL-4 antagonist	Antiasthmatics
Leukotriene synthase inhibitor	Antiasthmatics
Atrial peptidase inhibitor	Antihypertensive
Renin inhibitor	Antihypertensive
ACE inhibitor	Antihypertensive
Cholesterol acyltransferase	Cholesterol control
HMG-CoA reductase inhibitor	Cholesterol control
LDL receptor antagonist	Cholesterol control
Elastase inhiboitir	Cystic fibrosis
	Emphysema

While screening against targets at the protein level, for example, for inhibitors of enzymes and for molecules (agonists and antagonists) that interact with receptors, is expanding, the search for compounds that modulate gene expression, at the DNA level, is rapidly emerging.

BIOPHARMACEUTICALS

The term "biopharmaceutical" has been generally accepted for distinguishing the large-molecule pharmaceuticals (mainly proteins) that have emerged as a result of modern "bio-technology" from the traditional small-molecule drugs. The so-called bio-technology is largely based on two technological breakthroughs of the late 1970s: recombinant DNA and hydridoma methodologies. Over the past decade, genetic engineering or recombinant DNA technology has affected all facets of pure and applied biological research. The main thrust of the technology lies with the ability to introduce a manipulated genetic material into a cell in such a way as to permit its replication and passage to progeny cells. The primary, commercial application of genetic engineering has been to produce a protein of interest in large quantity by inserting its gene into a cell in which the protein can be efficiently expressed. The major steps involved in commercializing such a protein include:

1. Gene cloning using homologous probes, by direct expression, or through gene synthesis.
2. Construction of expression vectors.
3. Transforming of host cells.
4. Large-scale propagation of the transformed cell.
5. Recovery and purification of the cloned product.

The commonly used host cells for expressing heterologous genes include *E. coli*, yeast, and mammalian cells. The most commonly used

E. coli promoters are those from the *lac* or *trp* operons and the late promoter, *pL*, from phage lambda. These promoters can be turned on by the addition of an inducer (derepressor) such as IPTG (isopropyl thiogalactoside) or lactose for *lac*, or IAA (indoleacrylic acid) for *trp*, or by a temperature shift from 30 to 37°C to 42°C for *pL*. Following process optimization and scale-up, an expression level of 1 to 5 g/liter is achievable with *E. coli* expression cultures.

For yeast, there are in general two modes of transformation, integrative and autonomous. The integrative vectors are simple plasmids, and achieve clonal expression through integration into the genome by homologous recombination. The autonomous vectors possess a sequence known as ARS (autonomously replicating sequence). It, like an origin of replication, allows plasmids to be passed to progeny cells. The yeast 2 micron circle, a naturally occurring high-copy-number plasmid, is a commonly used ARS vector. The autonomous transformants are in general less stable than the integrative ones. The inclusion on an ARS plasmid of a yeast centromeric sequence can overcome the instability problem. To achieve efficient expression, several yeast promoters are used, PGK promoter from phosphoglycerate kinase, GAL1 from galactose kinase, PHO5 from acid phosphatase and HSP90 from a heat shock protein. Certain "supersecreting" hosts have been sought. An optimized host/vector yeast system can produce a heterologous protein up to 1 g/liter level, with 1 to 100 mg/liter more the norm.

Mammalian plasmid-based vectors, in addition to a prokaryotic replicon and a selection marker to permit DNA replication in *E. coli*, commonly have a eukaryotic replicon and a eukaryotic selection marker. The replicon usually comes from viruses such as Simian virus 40 (SV40), bovine papilloma virus (BPV) or Epstein Barr virus (EPV). The commonly used promoters are the SV40 early promoter (including its upstream enhancer region) and the metallothionine promoter (including its enhancer and metal regulatory elements). Some plasmid-based vectors also contain a splicing site to promote RNA stability and a poly-adenylation sequence to ensure transcription termination. The most commonly used mammalian hosts are CHO (Chinese hamster overy) and BHK (baby hamster kidney).

The selection of an expression system for a cloned gene depends on many factors: host/vector stability, expression level, ease of scale-up, post-translational modifications (glycosylation, secretion, amidation, folding, proteolysis, phosphorylation, etc.), product immunogenecity, and heterogenicity.

The biopharmaceuticals that have been approved for marketing include insulin (for diabetics), human growth hormone (for dwarfism), hepatitis B vaccine, alpha interferon (for hairy cell leukemia, Kaposi's sarcoma, genital warts and hepatitis C), gamma interferon (for chronic granulomatous disease), monoclonal antibody to treat transplant rejection, tissue plasminogen activator (for myocardial infarction, pulmonary embolism, and peripheral arterial occlusive disease), erythropoietin (for anemia caused by chronic renal failure), and colony stimulating factors (for leukopenia arising from cancer chemotherapy and for acceleration of engraftment in bone marrow transplant). Their combined sales in 1990 were about $2 billion. With expanding indications and markets of the approved biopharmaceuticals and with approvals of new products such as Factor VIII:C, Factor IX, monoclonal antibody for treating septic shock, IL-2, IL-4, epidermal growth factor, fibroblast growth factor, bone morphogenic proteins, HIV vaccines, superoxide dismutase, nerve growth factors, and cancer-suppressing factors, biopharmaceutical sales by the year 2000 are expected to exceed $10 billion. They will have a profound, positive impact on the maintenance and management of the health and harmony of the human body.

REFERENCES

1. Aiba, S., Humphrey, A. E., and Milis, N., *Biochemical Engineering*, 2nd ed., University of Tokyo Press, Tokyo, 1973.
2. Aida, K., Ichiba, I., Nakayama, K., Takinami, K., and Yamada, H. (Eds.), *Biotechnology of Amino Acid Production*, Elsevier, Amsterdam, 1986.
3. Anderson, S., Marks, C. B., Lazarus, R., Miller, J., Stafford, K., Seymour, J., Light, D., Rastetter, W., and Estell, D., *Science*, **230**, 144 (1985).
4. Aunstrup, K., "Enzymes of Industrial Interest—Traditional Products," *Ann. Report of Fermentation Processes*, **6**, 175 (1983).
5. Armiger, W. B. (Ed.), "Computer Applications in Fermentation Technology," *Biotechnology and Bioengineering Symposium Series No. 9*, John Wiley & Sons, New York, 1979.
6. Atkinson, B., and Mavituna, F., *Biochemical Engineering and Biotechnology Handbook*, The Nature Press, 1983.
7. Bailey, J. E., and Ollis, D. F., *Biochemical Engineering Fundamentals*, McGraw-Hill Book Company, New York, 1977.
8. Baumberg, S., Hunter, I., and Rhodes, M. (Eds.), *Microbial Products: New Approaches*, Cambridge University Press, Cambridge, England, 1989.
9. Baker, S. A., "Industrial Enzyme Process," *Sci. Prog. Oxf.*, **65**, 477 (1980).
10. Bull, A. T., Ellwood, D. C., and Ratledge, C. (Eds.), *Microbial Technology: Current State, Future Prospects*, Cambridge University Press, Cambridge, England, 1979.
11. *Chemical Marketing Reporter*, Schnell Publicating Company, New York.
12. Cheremisinoff, N. P., *Gasohol for Energy Production*, Ann Arbor Science Publishers Inc., Ann Arbor, MI, 1979.
13. deBoer, L., and Dijkhuizen, L., "Microbial and Enzymatic Processes for L-Phenylalanine Production." *Advances in Biochemical Engineering/Biotechnology*, **41**, 1 (1990).
14. Demain, A. L., "Riboflavin Oversynthesis," *Ann. Rev. Microbiol.*, **26**, 369 (1972).
15. DiMarco, A., and Pennella, P., "The Fermentation of the Tetracyclines," *Prog. Ind. Microbio.*, **1**, 45 (1959).
16. Esaka, N., Soda, K., Kumagai, H., and Yamada, H., "Recent Advances in Enzymatic Synthesis of Amino Acids," *Biotech. Bioeng.*, **22** (Suppl. 1), 127 (1980).
17. Friendmann, H. C., and Cagen, L. M., "Microbial Biosynthesis of B_{12}-like Compounds," *Ann. Rev. Microbiol.*, **24**, 159 (1970).
18. Godfrey, T., and Reichelt, J., *Industrial Enzymology*, The Nature Press, 1983.
19. Hatch, R., "Computer Applications for Analysis and Control of Fermentation," *Ann. Report of Fermentation Processes*, **5**, 291 (1982).
20. Hirose, Y., and Shibai, M., "Amino Acid Fermentation," *Biotech. Bioeng.*, **22** (Suppl. 1), 111 (1980).
21. Ho, N., "Yeast Alcohol Tolerance and Recombinant DNA for Improved Alcohol Production," *Ann. Report of Fermentation Processes*, **4**, 235 (1980).
22. Hopps, H. E., and Petricciani, J. C. (Eds.), *Abnormal Cells, New Products and Risk*, Tissue Culture Association, Gaithersburg, MD, 1985.
23. Humphrey, A. E., "Biochemical Reaction Engineering," Chapter 8 in *Chemical Reaction Engineering Reviews*, ACS Symposium Series 72, 1979.
24. International Technical Information Institute, "Japan's Most Advanced Industrial Fermentation: Technology and Industry," International Technical Information Institute, Tokyo, 1977.
25. Kalisz, H. M., "Microbial Proteinases," *Advances in Biochemical Engineering/Biotechnology*, **36**, 1 (1988).
26. Lowenhem, F. A., and Moran, M. K. (Eds.), *Faith, Keyes and Clark's Industrial Chemicals*, 4th ed., John Wiley & Sons, New York, 1975.
27. Magee, R. J., and Kosaric, N., "The Microbial Production of 2,3-Butanediol," *Advances in Applied Microbiology*, **32**, 89 (1987).
28. Martin, J. F., and Demain, A. L., "Control of Antibiotic Biosynthesis," *Microbial Rev.*, **44** (2), 230 (1980).
29. Martin, J. F., and Liras, P., "Enzymes Involved in Penicillin, Cephalosporin and Cephamycin Biosynthesis," *Advances in Biochemical Engineering/Biotechnology*, **39**, 153 (1989).
30. Miller, L., "Baculoviruses as Gene Expression Vectors," *Ann. Rev. Microbiol.*, **42**, 177 (1988).
31. Morris, C. E., "America's Gold," *Food Engineering*, **52** (5), 95 (1980).
32. Nakayama, K., "The Production of Amino Acids," *Process Biochemistry*, **11** (2), 4 (1976).
33. Nuesch, J., Heim, J., and Treichler, H. J., "The Biosynthesis of Sulfur-Containing Beta-Lactam Antibiotics," *Ann. Rev. Microbiol.*, **41**, 25 (1987).
34. Peppler, H. J., and Perlman, D. (Eds.), *Microbial Technology*, Vols. 1 and 2, 2nd ed., Academic Press, New York, 1979.
35. Parker, W. L., O'Sullivan, J., and Sykes, R., "Naturally Occurring Monobactams," *Advances in Applied Microbiology*, **31**, 181 (1986).
36. Perlman, D., "Influence of Penicillin Fermentation Technology to Processes for Production of Other Antibiotics," *Process Biochemistry*, **10** (9), 23 (1975).

37. Perlman, D., and Tsao, G. T. (Eds.), *Annual Reports on Fermentation Processes*, Vols. 1–3, Academic Press, New York, 1977–79.
38. Reed, G., and Peppler, H. J., *Yeast Technology*, AVI Publishing Company, Westport, CT, 1973.
39. Schierholt, J., "Fermentation Processes for the Production of Citric Acid, *Process Biochemistry*, **12** (9), 20 (1977).
40. Tabor, J. (Ed.), *Genetic Engineering Technology in Industrial Pharmacy*, Marcel Dekker, New York, 1989.
41. Tannenbaum, S., and Wang, D. I. C. (Eds.), *Single Cell Protein II*, M.I.T. Press, Cambridge, MA, 1975.
42. Tsao, G. T. (Ed.), *Annual Reports on Fermentation Processes*, Vols. 4–8, Academic Press, New York, 1980–85.
43. Weinstein, M. J., and Wagman, G. H. (Eds.), "Antibiotics: Isolation, Separation and Purification," *Journal of Chromatography*, **15** (Elsevier, Amsterdam, 1978).
44. U.S. International Trade Commission, *Synthetic Organic Chemicals: U.S. Production and Sales*, U.S. Government Printing Office, Washington, DC, 1989.
45. Woodruff, H. B., and Burg, R. W., "The Antibiotic Explosion," *Pharmacological Methods, Receptors and Chemotherapy*, Elsevier, Amsterdam, 1986.
46. Rushton, J. H. et al., *Chem. Eng. Prog.*, **46**, 467 (1959).
47. Miller, D. N., *A.I.Ch.E.J.*, **20**, 445 (1975).

25

The Pharmaceutical Industry

Jeffrey H. Watthey*

INTRODUCTION

The pharmaceutical industry is one of the major contributors to health care in the United States. The industry is almost exclusively concerned with the provision of prescription and over-the-counter drugs to meet human needs, but veterinary drugs also represent a significant market segment.

Government statistics indicate that costs to consumers for drugs and other medical nondurables amounted to $44.6 billion in 1989. This represented 7 percent of health-care costs nationwide, which in turn amounted to $604.1 billion or 11.6 percent of the gross national product (GNP) in 1989. Although health-care costs as a percentage of the GNP more than doubled from 1960 through 1989, the percentage of the total national health-care expenditure allocated to drugs decreased markedly, from 15.5 percent in 1960 to 7 percent in 1989.[1]

In 1989 approximately 13,000,000 people were involved in patient care in the United States, including 450,000 physicians. In the same year the pharmaceutical industry employed 185,900 people, 22 percent of whom were involved in medical research and development, according to PMA statistics.[8]

In 1990 there were over 650 companies in the United States that manufactured drugs, but the top 10 U.S. companies accounted for 86.5 percent of domestic drug sales. Generic prescription drugs are estimated to be about 20 percent of the total market. Their use is emphasized in hospitals, which account for 20 percent, and health maintenance organizations (HMOs), which are responsible for 15 percent of prescription purchases. Mail order pharmacies dispense an estimated 5 percent of all outpatient prescriptions. In addition, shipments of nonprescription drugs reached an estimated $10.4 billion in 1990. Research expenditures accounted for 16.85 percent of pharmaceutical sales.

Worldwide sales for 1988 are estimated by Robert Fleming Securities to have been $128 billion.

DEVELOPMENT OF THE INDUSTRY

The products marketed by the pharmaceutical industry range from a few drugs that have

*CIBA-GEIGY Pharmaceuticals Division.

been known for over one hundred years to recent introductions consequent on the latest scientific and technical advances.

The industry has assumed the responsibility for providing efficacious and safe drugs for the treatment of disease, and an array of government controls have been imposed on drug manufacturers to ensure that this responsibility is met.

Drugs that were available before the dawn of the era of commercial drug development were derived from folk medicine and as a group are largely discredited today. However, there are notable exceptions such as quinine, morphine, and digitalis. Despite a current preoccupation with innovation that is driven by scientific and technological advances as well as commercial considerations, there are a number of drugs of proven efficacy that have been available for over 50 years. Some prominent examples are given in Table 25.1.

The discovery of the antibacterial activity of sulfanilamide in 1932 and that of penicillin in 1940 were pivotal events that served to establish the drug industry in its present form. An intensive research effort was initiated that continues at an ever increasing pace and has made available many new and effective drugs. These drugs form the basis of current nonsurgical treatment of disease. Some examples of these drugs are given in Table 25.2.

The cost of developing a new drug is increasing steadily, because of the increasing complexity of modern research as well as the costs of meeting regulatory requirements.

TABLE 25.1 Some Older Drugs (Over 50 Years) of Proven Efficacy

Drug	*Indication*	*Year of Introduction*
Quinine	Antimalarial	1639
Smallpox vaccine		ca. 1800
Morphine	Analgesic	1820
Digitalis	Cardiac stimulant	1785
Aspirin	Analgesic	1894
Phenobarbital	Hypnotic	1912
Insulin	Hypoglycemic	1921
Sulfanilamide	Antibacterial	1932

TABLE 25.2 Examples of More Recent Effective Drugs

Class	*Name*	*Year of NDA Approval*
Analgesic	Meperidine	1944
	Pentazocine	1967
Anti-allergy	Cromolyn Sodium	1973
Antibacterial	Sulfisoxazole	1949
	Cotrimoxazole	1973
Antibiotic	Penicillin	1940
	Chlortetracycline	1948
	Erythromycin	1952
	Cephalexin	1971
Anticonvulsant	Carbamazepine	1968
Antidepressant	Imipramine	1959
	Amitriptyline	1962
	Fluoxetine	1987
	Clomipramine	1989
Antihistamine	Astemizole	1988
Antihypertensive	Prazosin	1976
	Metoprolol	1978
	Captopril	1981
	Nifedipine	1981
	Enalapril	1984
Anti-inflammatory	Cortisone Acetate	1950
(Steroidal)	Dexamethasone	1958
(Nonsteroidal)	Indomethacin	1965
	Ibuprofen	1974
	Naproxen	1980
	Diclofenac Sodium	1988
Antineoplastic	Fluorouracil	1962
	Cisplatin	1978
Antipsychotic	Chlorpromazine	1954
	Haloperidol	1967
	Clozapine	1990
Anti-ulcer	Cimetidine	1977
	Ranitidine	1983
Antiviral (AIDS)	Zidovudine	1987
Anxiolytic	Chlordiazepoxide	1960
	Diazepam	1963
	Alprazolam	1981
	Buspirone	1985
Cardiotonic	Amrinone	1983
Diuretic	Chlorothiazide	1958
	Hydrochlorothiazide	1959
	Spironolactone	1960
	Furosemide	1966
Estrogen	Estradiol	1940
Hypocholesteremic	Gemfibrozil	1981
	Lovastatin	1987
Hypoglycemic	Tolbutamide	1957
	Chlorpropamide	1958
Immunosuppressive	Cyclosporine	1983
Vaccine	Poliomyelitis	1961
	Measles	1964
	Rubella	1967
	Hepatitis B	1986

Estimates from a recent study by the U.S. Center for Drug Development at Tufts University[2] place the cost of research and development for a new chemical entity at $231 million in 1987 dollars, and the time required to bring a product to market at 12 years. Only five of 4000 compounds screened in preclinical testing survive to human testing, and only 20 percent of these obtain FDA approval.

Although the drug industry involves the application of scientific research and technology to combating disease, it is at the same time a highly successful commercial endeavor. Table 25.3 lists the 25 top-selling branded drugs worldwide, divided into categories and with 1990 estimates of sales figures.

The drug discovery and development process is not only increasingly more costly and more time consuming; its fundamental nature is changing.

In the process that had been standard in the past, the initial "screening" involved evaluation of novel substances prepared by synthetic organic chemists and designed to have the potential for biological activity. These tests, usually in intact animals, were

TABLE 25.3 Top-Selling Branded Drugs Worldwide, with 1990 Estimates of Sales Figures

			$B
	Anti-ulcer Agents		
1	Ranitidine (Glaxo/Sankyo)	H-2 antagonist	2.78
6	Cimetidine (SmithKline Beecham)	—	1.10
22	Famotidine (Merck & Co.)	—	0.50
	Antihypertensives		
2	Nifedipine (Bayer/Takeda & Pfizer)	Ca antagonist	1.95
3	Captopril (Bristol-Myers Squibb)	ACE inhibitor	1.50
4	Enalapril (Merck & Co.)	—	1.48
5	Atenolol (ICI)	Beta blocker	1.13
13	Diltiazem (Marion-Merrell Dow)	Ca antagonist	0.75
	Nonsteroidal Anti-inflammatory Agents		
7	Diclofenac Na (CIBA-GEIGY)		1.04
16	Naproxen (Syntex)		0.69
18	Piroxicam (Pfizer)		0.64
	Antibiotics		
8	Cefaclor (Eli Lilly)	Oral cephalosporin	0.84
9	Ciprofloxacin (Bayer)	Quinolone antibiotic	0.80
14	Clavulanate K (SmithKline Beecham)	Combination antibiotic	0.71
17	Ceftriaxone (Hoffmann-La Roche)	—	0.67
23	Cefotaxime Na (Hoechst-Roussel)	Antibiotic	0.49
24	Amoxicillin (SmithKline Beecham)	Antibacterial	0.44
	Miscellaneous		
10	Albuterol (Glaxo)	Bronchodilator	0.78
11	Fluoxetine (Eli Lilly)	Antidepressant	0.77
12	Lovastatin (Merck & Co.)	Hypolipemic	0.75
15	Acyclovir (Wellcome)	Antiviral	0.70
19	Iopamidol (Schering AG)	Contrast medium	0.60
20	Iohexol (Sterling/Daiichi/Schering AG/Nycomed)	—	0.55
21	Terfenadine (Marion-Merrell Dow)	Anti-allergy	0.52
25	Alprazolam (UpJohn)	Anxiolytic	0.44

*Barclays de Zoete Wedd, "Pharmaceutical Industry Perspectives."[3]

designed to assess any activity of interest that the substance might have. When desirable biological activity was encountered, the active substance would be systematically modified until it was thought that the best possible analogue had been obtained. Secondary tests and safety evaluation followed.

The new process, which has now become dominant, derives from the realization that a new commercially successful drug has to be conceived from the outset as meeting a significant medical need. Once a broad target area has been defined (e.g., arthritis, hypertension, schizophrenia), a search is made for actual or potential specific biochemical processes that, if modulated, would be predicted to have a beneficial fundamental effect on the course of the disease. Once the biochemical process has been defined, an *in vitro* assay is developed, so that the effect of any given compound on that process can be determined. With as much logic as can be mustered, candidate compounds are designed, then synthesized, and evaluated. Once the desired biochemical activity has been obtained, the active substances are evaluated in an *in vivo* model to establish that oral administration of the test substance results in the desired effect in an intact animal species. Should this not be the case, poor absorption and/or rapid formation of inactive metabolites may be to blame, and appropriate modification of the lead compounds will be undertaken.

The biochemical process that is modulated usually is either the enzymatic transformation of a key intermediate in the synthesis of a hormone or a neurotransmitter or, alternatively, the interaction of such a substance with a receptor protein, which would normally elicit a signal resulting in the triggering of a specific biochemical event.

This change to a more focused, logical strategy is philosophically most profound, and future developments in the area of computer-assisted molecular modeling will render it even more so. However, its impact is being felt today for two more pragmatic reasons. The first of these is a significant reduction in the number of animals used for the evaluation of potential drug substances, which meets the challenge posed by animal welfare concerns, but also is motivated by considerations of the reliability of the experimental data as well as cost. The second is the emergence of a newly defined challenge: bioavailability was guaranteed when the initial test result came from an *in vivo* assay; but today the determination of activity is based upon *in vitro* screening results, and significant developmental efforts may be required to provide a candidate drug with adequate bioavailability. It thus has become necessary that bioavailability and drug metabolism studies (designed, respectively, to investigate (1) the levels of the candidate drug in plasma and various organs and tissues following differing dosing regimens, and (2) the manner in which the administered test substance is chemically transformed by enzyme systems within the body) be conducted much earlier in the drug development process. This approach implies that more such studies must be undertaken, adding to the cost of the drug development process. One consequence of this is the realization that the physical properties of the candidate drug should come within certain values if optimum biological activity is to result. To achieve this goal, it sometimes is necessary to prepare a "pro-drug"—a substance that is well absorbed and biologically inactive, but is converted *in vivo* by a metabolic enzyme system into a biologically active drug which itself would not be well absorbed and would not be a satisfactory drug if administered directly.

DRUG REGULATION AND PATENTS

The driving force in contemporary drug research is the development of new products that will be covered by worldwide patents and will meet the requirements of the regulatory agencies in the various countries in which they will be marketed. Today, major drug companies are either multinational organizations or have created strategic marketing alliances that will enable any significant new product to be marketed throughout the world. This strategy offers the maximum return on the

considerable investment required in the research and development process.

The first step in what might be called the regulatory process involves the application for a patent, usually in the country in which the drug substance was first discovered, and subsequently worldwide. In the United States, in a contested situation, the applicant who can demonstrate that he or she was the first inventor will prevail, whereas in most other countries worldwide the patent will be granted to the first to file the application. International conventions provide that patents may be filed outside the country where the first patent was filed within one year of the initial filing date. A U.S. patent is valid for 17 years from the date of issue. However, competitive pressures often will result in relatively early filing with consequent issue of the patent before the drug is approved, so a portion of this 17-year period will have passed before the drug can be marketed. This problem was addressed by the Drug Price Competition and Patent Term Restoration Act of 1984, also known as the Waxman–Hatch Act. If all requirements are met, the life of the patent may be extended for up to five years by one-half of the period between the filing of the IND (see below) and the filing of the NDA plus all the time between the NDA filing and the NDA approval, to a maximum patent life of 14 years after the first marketing approval. The Waxman–Hatch Act also provides for a 5 year market exclusivity for the first NDA approval of a new drug pertaining to a particular patent.

The regulation of the distribution of drugs to the public dates from February 1, 1906, when the Pure Food and Drug Act was enacted. It was recognized from the start that the provisions of the act were inadequate, but it was not until 1938 that the Food, Drug, and Cosmetic Act was signed into law. This act outlawed drugs that were dangerous when used as directed, provisions relating to formula disclosure were established, and a requirement was established that made it mandatory for the manufacturer of a drug to demonstrate to the satisfaction of the Food and Drug Administration (FDA) that it was safe. In 1962, the Kefauver–Harris amendments to the Food, Drug, and Cosmetic Act were adopted. These amendments gave the FDA regulatory authority over the assessment of the efficacy as well as the safety of new drugs. The driving force for this legislation was the incidence in a number of European countries of phocomelia (shortened, malformed limbs) in newborn children whose mothers had taken thalidomide, a sedative-hypnotic, during pregnancy. The drug never was approved for use in the United States.

Precise regulations govern the investigation of new drugs. At least 30 days before clinical investigation of a new drug can commence, the Notice of Claimed Investigational Exemption, the so-called IND, must be filed with the FDA. This document must include detailed descriptions of in-depth laboratory studies of the drug's biological activity and safety in animals, a description of the structure and synthesis of the new compound, and the formulation to be used for clinical studies. The 1962 regulations prescribe that clinical trials be conducted in three phases. In Phase I, tolerance, absorption, excretion, and metabolism are studied in normal (i.e., well) volunteers. These studies are conducted at the dose ranges anticipated for the trials in patients, and are intended to establish that these doses may be administered safely. Phase II covers initial trials in a limited number of patients with the goal of determining efficacy. Phase III studies provide an assessment of the drug's safety and efficacy in a large number of patients. Usually these studies involve comparison with a preexisting drug, if that is appropriate, and also with a placebo in order to verify efficacy. Before any of these clinical studies can be initiated, the protocols must be reviewed and approved by the FDA. When sufficient studies have been completed, all results and related information are assembled and submitted to the FDA as the New Drug Application (NDA). The NDA is a truly voluminous document, and FDA review easily can take two years. FDA approval was granted to 23 New Chemical Entities in both 1989 and 1990, and 30 efficacy supplements also were approved in each of those years. The latter are submissions of already approved

drugs for additional uses, doses, or dosage forms, which require additional clinical trials by the sponsor and subsequent review by the FDA.

There are also federal regulations governing the preclinical studies (specifically the animal safety or toxicology studies) referred to above—the Good Laboratory Practice (GLP) regulations, which became effective in 1979. These regulations delineate the procedures that must be followed during the conduct of animal safety studies, including a precise protocol, the designation of a study director for each study, and a quality assurance unit to inspect and audit each study and facility and to assure tope management that the study has been conducted in compliance with the GLPs, and so on. These regulations have had a positive impact on the quality of the safety studies, although at a significant increase in cost.

The Kefauver–Harris Amendments of 1962 also required that methods, facilities, and controls used by the manufacturer conform with "current good manufacturing practice." The FDA issued regulations (GMPs) in 1962 defining current good manufacturing practice. These regulations were revised in 1971 and again in 1979. The newer revision includes specific requirements for the establishment of a quality control unit and written procedures (Standard Operating Procedures or SOPs) in certain key areas. Also mandated are expiration dates for essentially all marketed drug products. The FDA moniters compliance with these regulations by means of periodic unannounced inspections.

PRODUCTS

Representative currently manufactured drugs from the various categories listed in Table 25.3 are described below. The chemical structures provide an indication of the structural diversity present in drug active ingredients. More detailed descriptions of the various agents may be found in *Goodman and Gilman's The Pharmacological Basis of Therapeutics.*[4]

Analgesics (Narcotic)

These substances are used primarily for the relief of severe pain. They all have physical dependence liability and the potential to cause respiratory depression.

Morphine sulfate

Pentazocine

Propoxyphene Hydrochloride

Analgesics (Nonnarcotic)

These substances are widely used nonprescription drugs. They find application for the relief of mild pain but also are used for the relief of symptoms of rheumatoid arthritis.

Aspirin

Acetaminophen

Phenacetin

Ibuprofen

Sulfisoxazole

Trimethroprim

Sulfamethoxazole

Anti-allergy Agents and Bronchodilators

Cromolyn sodium is administered by inhalation. Its mode of action involves the inhibition of the release of histamine and other inflammatory mediators from several cell types. Its main use is for the prophylactic treatment of bronchial asthma. Inhalation of a selective [beta-2]-adrenergic agonist such as albuterol is the preferred form of bronchodilator therapy for asthma.

Cromolyn Sodium

Albuterol

Antibacterials

Before penicillin and the other antibiotics became generally available, the sulfonamides were the mainstay of antibacterial chemotherapy. They maintain a significant role, largely due to the use of a combination of sulfamethoxazole and trimethoprim (co-trimoxazole) for the treatment of bacterial respiratory tract infections and gastrointestinal infections.

Antibiotics

The term "antibiotics" originally was restricted to chemical substances produced by certain species of microorganisms that suppressed the growth of other microorganisms with the possibility of eventually destroying them. These substances were applied very effectively to the control of human bacterial infections. The term now includes partially or totally synthetic variants with activity against wider ranges of bacteria (e.g., cefaclor) and/or resistant strains (e.g., ceftrioxone). The introduction of these newer agents has reduced the applicability of some of the original antibiotics, and erythromycin, streptomycin, and chlortetracycline are no longer the major forces they once were. However, the penicillins remain first-line agents for the treatment of a variety of major infectious diseases. Clavulanic acid inhibits [β]-lactamase, an enzyme capable of destroying penicillins and cephalosporins. Imipenem and cilastatin are components of the combination product Primaxin. Cilastatin inhibits a dipeptidase that is capable of rapidly hydrolyzing imipenem.

Erythromycin

Chlortetracycline Hydrochloride

Penicillin G Sodium

Cephalexin

Ampicillin

Cefaclor

Streptomycin Sulfate

Ceftriaxone Sodium

Clavulanate Potassium

Cilastatin Sodium

Imipramine Hydrochloride

Imipenem

Amitriptyline Hydrochloride

Anticonvulsants

These agents are used for the treatment of the epilepsies. The sedative effects of phenobarbital and its tendency to disturb behavior in children have reduced its use as a primary agent.

Phenobarbital

Phenytoin

Carbamazepine

Valproic Acid

Antidepressants

The first antidepressant, imipramine, was introduced in 1959. The newer agents such as fluoxetine have fewer side-effects than the classical tricyclics. Clomipramine is used for the treatment of obsessive compulsive disorder.

Fluoxetine

Clomipramine Hydrochloride

Antihistamines

The agents in this category are [H-1] antagonists, as opposed to the [H-2] antagonists which are discussed under anti-ulcer agents. The major application of [H-1] antagonists is in the symptomatic relief of various diseases of allergy such as hay fever. The newer agents, astemizole and terfenadine, lack the sedative side-effects traditionally associated with this group.

Diphenhydramine Hydrochloride

Chlorpheniramine Maleate

Tripelennamine Hydrochloride

Astemizole

Terfenadine

Antihypertensives

In addition to the diuretics, which are listed separately, first-line antihypertensive drugs may be grouped into three categories: sympatholytic agents ([β] and [α] blockers, e.g., metoprolol and prazosin, respectively), angiotensin converting enzyme inhibitors (e.g., captopril, enalapril), and calcium channel blockers (e.g., diltiazem, nifedipine). These agents all have a very low incidence of side-effects, and treatment is being extended to more patients with relatively moderate hypertension. In addition, therapy is becoming more individualized, with the choice of drug being determined by such factors as the patient's age, race, concomitant diseases and therapies, and life-style. Drugs in all three categories have application in the treatment of congestive heart failure.

Diltiazem Hydrochloride

Prazosin Hydrochloride

Metoprolol Tartrate

Captopril

Nifedipine

Enalapril Maleate

Anti-inflammatories (Steroidal)

These drugs are reserved for situations in which other therapy has proved unsuccessful, for there is a significant potential for serious complications.

Cortisone Acetate

Dexamethasone

Anti-inflammatories (Nonsteroidal)

Drugs such as naproxen, diclofenac sodium, and piroxicam, in addition to the nonnarcotic analgesics discussed above, are used extensively for the symptomatic treatment of such conditions as rheumatoid arthritis and osteoarthritis. Considerable efforts are being made by the industry to provide generally acceptable agents that will modify the course of these diseases. Of the currently available disease-modifying drugs, gold compounds such as auranofin are able to induce a protracted remission in about 15 percent of patients. The toxic effects are significant.

Naproxen

Diclofenac Sodium

Piroxicam

Auranofin

Antineoplastics

Fundamental progress continues to be made in cancer chemotherapy. This progress is not characterized by the adoption of major new drugs, but by the improved design of regimens for the administration of the available drugs, and by a greater understanding of the mechanism of action of antitumor agents and of the resistance to them. Drugs having several different mechanisms of action are available, such as alkylating agents (e.g., cyclophosphamide), antimetabolites (e.g., fluorouracil and methotrexate), and platinum coordination complexes such as cisplatin.

Fluorouracil

Cisplatin

Cyclophosphamide

Methotrexate

Antipsychotics

These drugs are used for the management of patients with psychotic or other serious psychiatric illness. The drugs are regarded as safe, but neurological side-effects (muscle spasms, restlessness, parkinsonism) are commonplace. Clozapine is superior in this regard, but with this drug the incidence of agranulocytosis approaches 1 percent.

Chlorpromazine

Haloperidol

Thioridazine

Clozapine

Anti-ulcer Agents

These agents are inhibitors of gastric acid secretion (histamine [H-2] antagonists) and offer very effective treatment of peptic ulcers.

Cimetidine

Ranitidine Hydrochloride

Antiviral Agents

Development of effective antiviral agents has been slow, as the drugs must be highly selective because viral replication depends largely on the metabolic processes ongoing in the invaded cell. Zidovudine is the most important agent available for the treatment of AIDS, and Acyclovir is effective for the treatment of herpes simplex virus infections.

Zidovudine

Acyclovir

Anxiolytics

These drugs are widely used for the treatment of mild anxiety and tension but are not indicated for the treatment of psychoses.

Chlordiazepoxide

Diazepam

Alprazolam

Buspirone Hydrochloride

Cardiovascular Agents

Amrinone has been approved for use in the United States for the short-term management of congestive heart failure. The major use of digitoxin is in the treatment of chronic mild heart failure, frequently in combination with an angiotensin converting enzyme (ACE) inhibitor. This combination also is effective in the treatment of the acute condition. Nitroglycerin is representative of the nitrates used for the treatment of angina pectoris.

Amrinone

Digitoxin

Nitroglycerin

Cold Preparations

These medications are used to treat nasal and sinus congestion. Phenylpropanolamine is a vasoconstrictor, and usually is combined with an [H-1] antihistamine for this application.

Phenylpropanolamine Hydrochloride

Cough Preparations

Dextromethorphan is a widely used non-prescription antitussive agent.

Dextromethorphan Hydrobromide

Diuretics

Diuretics are used to increase the volume of urine excreted by the kidneys. They are effective in the treatment of edema associated with congestive heart failure. The thiazide diuretics (e.g., hydrochlorothiazide) have antihypertensive activity and are used to control blood pressure either alone or in combination with other antihypertensive agents. Spironolactone (an aldosterone antagonist) and amiloride are "potassium-sparing" diuretics, which cause retention of potassium unlike the thiazide diuretics, which promote potassium excretion. Combinations of hydrochlorothiazide with spironolactone and with amiloride are widely prescribed. Furosemide is an example of a "high-ceiling diuretic"; this class of agents elicits a peak diuretic effect far greater than that of the other types.

Hydrochlorothiazide

Spironolactone

Furosemide

Amiloride Hydrochloride

Estrogens

A major application is postmenopausal estrogen replacement therapy. Synthetic estrogens also are used as a component of oral contraceptive combination products (see below).

Estradiol

Hypocholesteremics

Lovastatin is indicated as first-line therapy for patients who are at high risk of myocardial infarction due to hypercholesterolemia. Gemfibrozil also elicits beneficial effects on plasma lipids and lipoproteins in hypercholesterolemia.

Gemfibrozil

Lovastatin

Hypoglycemics

These drugs are indicated for the treatment of non-insulin-dependent diabetes. Glipizide is representative of a more potent second generation of this category.

Tolbutamide

Chlorpropamide

Glipizide

Sedatives and Hypnotics

Although these drugs are employed in other situations, flurazepam being a benzodiazepine anxiolytic (see above), the indicated substances find particular application in this area.

Phenobarbital

Flurazepam Hydrochloride

Steroidal Oral Contraceptives

The most common type of oral contraceptive is a combination of a synthetic estrogen (such as mestranol) and a progestin (such as norethynodrel).

Mestranol

Norethynodrel

PRODUCTION

Chemical Manufacturing

A typical drug as administered to a patient may be in the form of a tablet, a capsule, or a solution for injection; and it contains not only the active ingredient, but also additional substances that serve to promote the stability of the drug and aid in its reproducible dissolution and/or distribution. This section addresses the manufacture of the active ingredients.

Most drug active ingredients are prepared by chemical synthetic, generally by a batch process in which a specific quantity of a starting material is converted by successive chemical reactions to the final product. Antibiotics are an exception, being obtained by a fermentation process. The product is obtained by extraction, and may be chemically modified using standard synthetic procedures. In recent years drug products derived from biotechnology processes have been introduced to the marketplace. Using genetic engineering techniques, common bacteria may be induced to prepare important peptide drugs normally produced by a totally different organism—the human. In this way, drugs such as human insulin have become available.

In designing a synthetic sequence, consideration must be given to the cost and availability of starting materials and their toxicity, as well as the toxicity of the intermediates and possible by-products. The disposal of waste material from the synthetic processes is of major concern, as it is the subject of federal, state, and local regulations. Consequently, every effort is made to maximize the recycling of solvents, and so on, and to obtain as much pure product as possible from the process runs. Reactions generally are carried out in reactors varying in size from 50 to several thousand gallons. Depending on the reaction, either stainless steel or glass-lined reactors are used. Figure 25.1 shows a drawing of a typical 100-gallon jacketed reactor equipped with an agitator, various inlets, and a bottom outlet. Figure 25.2 shows a series of reactors installed in a production facility for Geopen (R), a semisynthetic penicillin derivative.

The following paragraphs describe processes typical of those used in the preparation of pharmaceutical active ingredients.

Acetylsalicylic Acid. A flow diagram for a commercial process for producing aspirin is shown in Fig. 25.3. A reactor is charged with salicylic acid powder, acetic anhydride (the acetylating agent, a liquid), and the "mother liquor" from a previous process run (see below). The mixture is stirred for 2 to 3 hours and then transferred to a crystallizing kettle where the temperature is reduced to 3°C to promote crystallization of the product. A portion of the supernatant liquid (mother liquor) is drawn off for use in the next run, and the remaining slurry is centrifuged. The resulting crystalline material is dried in a rotary drier to yield bulk aspirin. Excess mother liquor is distilled to yield acetic anhydride, which is recycled, and acetic acid.

DIMENSIONS, IN.

25 PSI VESSELS

Capacity, Gallons	A	B	C	D	E Seal	E St. Box
100	36	41¼	3	67½	97½	99⅞

Fig. **25.1**. Plan and elevation of a 100-gallon glass-lined reactor. (*Courtesy Pfaudler Co.*).

Tripelennamine. A flow diagram for the production of tripelennamine (Pyribenzamine®) is shown in Fig. 25.4.

Step 1. A reactor is charged with thionyl chloride and toluene, and a solution of dimethylaminoethanol in toluene is added at a controlled rate. The reaction mixture then is heated at the boiling point to complete the reaction. Sulfur dioxide (a gas) is evolved and absorbed in the scrubber. The reaction mixture is cooled to promote crystallization

Fig. 25.2. Production facilities for Geopen. (*Courtesy Pfizer, Inc.*).

of the intermediate product, 2-chloro-*N*,*N*-dimethylethylamine hydrochloride, which is separated by centrifuging.

Step 2. A reactor is charged with 2-aminopyridine, benzaldehyde, and formic acid. The mixture is heated at the boiling point until the reaction is complete. After cooling, the reaction mixture is acidified by the addition of aqueous hydrochloric acid. Toluene is added to the reactor, and the mixture is stirred to transfer the nonbasic material to the upper, nonaqueous layer. The lower, acidic layer is transferred via a bottom outlet to a new reactor. Excess 30 percent aqueous sodium hydroxide is added; the intermediate 2-benzylaminopyridine separates out and is removed by centrifuging.

Step 3. A reactor is charged with water and 2-chloro-*N*,*N*-dimethylethylamine hydrochloride. To the resulting solution is added an equivalent amount of 30 percent aqueous sodium hydroxide. Toluene is added to the reactor, and the mixture is stirred to transfer the free base to the nonaqueous layer. The lower aqueous layer is drawn off, and the toluene solution of the free base is dried by the addition of anhydrous sodium carbonate and sodium hydroxide flakes.

Step 4. A reactor is charged with sodium amide and toluene. A toluene solution of the

Fig. 25.3. Flow diagram for production of bulk aspirin. (*Excerted by permission from* Chem. Eng., ***60*** *(6), 116–120, 1953; copyright 1953 by McGraw-Hill, Inc.*)

Fig. 25.4. Schematic process for tripelennamine.

2-benzylaminopyridine obtained in Step 2 is added at a controlled rate. After formation of the sodium salt of the 2-benzylaminopyridine is complete, a toluene solution of the 2-chloroethyl-*N*,*N*-dimethylethylamine from Step 3 is added slowly. After completion of the reaction, the reaction mixture is passed through a filter to remove the sodium chloride by-product. The filtrate is concentrated under reduced pressure, and the residual oily product is distilled under reduced pressure to yield tripelennamine base.

Step 2:

Step 1:

Step 3:

Step 4:

Penicillin. The production of an antibiotic is typified by the manufacture of penicillin. Originally, the means of production was the surface process in which the mold grows on the surface of a shallow layer of medium in trays or bottles. With the development of the commercial method of submerged fermentation, a reduction in space and labor requirements resulted in a very significant decrease in cost.

For the production of penicillin by submerged fermentation, the "seed" for the large fermentation tanks of 5000- to 30,000-gallon capacity is prepared by growing a master stock culture of *Penicillium chrysogenum* from lyophilized spores by incubation on a nutrient agar substratum. Several gallons of culture medium will be used to "seed" a large tank.

The four main stages in the manufacture of penicillin are:

1. Fermentation.
2. Removal of the vegetative matter from the fermentation broth and extraction of the antibiotic with solvents.
3. Purification and preparation of the appropriate stable salt derivative.
4. Assay.

The fermentation broth is prepared from corn steep liquor, to which 2 to 4 percent lactose is added, in addition to such inorganic materials as calcium carbonate, potassium dihydrogen phosphate, magnesium sulfate, and traces of iron, copper, and zinc salts. The pH of the medium is adjusted to 4.5 to 5.0, and then it is fed into the fermentor, which is equipped with an agitator, a means of introducing sterilized air, and a temperature control system. Following sterilization of the fermentor, the mold is introduced through sterile lines by air pressure. The temperature is kept between 23°C and 25°C while the mold is growing. Sterile air is supplied constantly, and agitation distributes the mold uniformly. Assays are made periodically, and after 50 to 90 hours when the potency no longer is increasing, the mold is harvested. The contents of the fermentor are cooled to 5°C because of the instability of penicillin at room temperature, and are filtered.

The filtered liquor is acidified to a pH of 2.5 with phosphoric acid and then subjected to continuous countercurrent extraction with amyl acetate and chloroform.

Influenza Vaccine. Influenza vaccines are sterile, aqueous suspensions of two or more strains of suitably inactivated influenza virus. The strains used in any particular situation are designated by the Division of Biologics Standards of the National Institutes of Health.

The production sequence starts with fertile chicken eggs that are disinfected on the exterior. A small hole is drilled through the shell, the virus is introduced, and the hole is sealed with gelatin. The eggs are incubated at 99°F for 48 hours, after which a circular top section of the egg is removed. The egg fluid is extracted, and the live virus is separated from the fluid by centrifuging. The relatively dense virus particles separate and accumulate on the centrifuge cartridge. The virus is collected and reconstituted in a saline solution, and then inactivated by treatment with formalin under defined conditions. The preparation is checked for virus inactivation, sterility, and potency (as an antigen), and is transferred into vials for distribution.

Pharmaceutical Production

Drugs seldom are administered as pure chemical compounds but generally are given in a formulation containing inert substances in addition to the active ingredient. It is the

function of the formulation pharmacist to devise a dosage form of the drug having the high degree of uniformity, physiological availability, and therapeutic quality expected of modern drug products. Prime considerations are the stability of the active ingredient (good shelf life) and its bioavailability. The many variables influencing these factors must be given adequate consideration, and the resulting formulation will require careful selection and quality control of the various ingredients employed. Pharmaceutical products also should be appealing to the patient; they should look pleasant and, if not flavored, should be free of objectionable odor and taste.

Drugs for specific applications are formulated into various types of dosage forms, which may be classified as follows:

- Tablets and capsules.
- Solutions, emulsions, and suspensions.
- Ointments, creams, and suppositories.
- Aerosols.
- Ophthalmic solutions.
- Parenteral solutions.
- Transdermal drug delivery systems.
- Advanced drug delivery systems.

A detailed discussion of the preparation of these dosage forms and other minor categories is presented in *Remington's Pharmaceutical Sciences.*[5] A brief overview of the production of tablets, capsules, and parenteral solutions is given below.

Tablets and Capsules

Tablets and capsules, the most frequently used dosage forms for oral administration, are greatly favored because they provide exact drug doses, protect the drug during storage, and are convenient and easy to take. However, before the drug contained in a tablet can be absorbed, the tablet, after ingestion, must disintegrate, and the drug must go into solution. Whatever impedes any step in the dissolution process can potentially slow drug absorption and thus decrease the bioavailability of the drug contained in the tablet; so tablets and capsules must be carefully and properly designed and manufactured.

Capsules present similar problems. The gelatin shell of the capsule must first dissolve in the stomach fluids to release the contents, which in turn must be dispersed and dissolved before absorption can occur.

Many factors influence dissolution rates, such as particle size, crystalline form, and the surface area of the drug per unit weight; diluents, disintegrants, and wetting agents in the formulation; and compressive force used in the processing of tablets. An estimate of the dissolution rate is provided by a determination of the time required to dissolve 90 percent of the drug contents of a tablet or capsule at $37.0 \pm 0.5°C$ under standard conditions. The results may be validated by bioavailability studies in humans. The relationships between elapsed time and blood levels attained following administration of the solid dosage form and an aqueous solution of the drug are compared. Ideally, identical blood levels should be reached, and the areas under the two time/blood level curves should be comparable.

In addition to the active ingredient, a tablet usually will contain a binder to provide cohesiveness, a filler to provide an acceptable-size tablet, a disintegrant that causes the tablet to break up when exposed to an aqueous medium, and a lubricant to prevent sticking in the tableting machine.

There are three general methods of tablet preparation: wet granulation, dry granulation, and direct compression. The purpose of the granulation procedure is to prepare a homogeneous mixture of the ingredients in a form that will flow freely in the processing machinery. A modern high-speed tableting machine capable of producing up to 10,000 tablets per minute is shown in Fig. 25.5.

In addition to uncoated compound tablets, the types listed below are produced for specific purposes:

- Sugar-coated tablets.
- Film-coated tablets.
- Enteric-coated tablets.
- Multiple-compressed tablets, layered tablets, and press-coated tablets.
- Prolonged-action tablets.

Fig. 25.5. Stokes GTP™ tableting press. (*Courtesy of Stokes, Pennwalt Corp.*)

Capsules are solid dosage forms in which the drug substance is enclosed in either a hard or a soft soluble shell of gelatin. The hard gelatin capsule consists of two sections, one slipping over the other; they are filled with the drug, usually mixed with a diluent and a lubricant to provide complete filling of the capsule and to ensure proper flow through the encapsulating machine. The filled capsules are sealed either by spotwelding or by applying a band of gelatin around the joint. Figure 25.6 illustrates the filling of hard gelatin capsules.

Soft elastic gelatin capsules are prepared in a rotary die process from two continuous gelatin ribbons. The capsules are formed, filled, and sealed in accurately timed operations, and can be filled with any liquid, semiliquid, paste, or even powder that will not dissolve or interact with the gelatin.

Parenteral Preparations

Parenteral preparations are administered by injection and can provide immediate physiological action. They are of special utility in those cases where the patient in unconscious. Also the therapeutic response of a drug is more readily controlled by parenteral administration than by oral dosages because the irregularities of intestinal absorption are avoided. However, because the protective barriers of skin and mucous membranes are bypassed, it is imperative that parenteral solutions meet the highest possible standards of purity and freedom from bacterial or viral contamination.

Single- and multiple-dose vials and ampules are used in medical practice. The vehicle (solvent), which is of greatest importance for parenteral products, is specially distilled, pyrogen-free water. The container of choice is glass.

Prior to large-scale production, painstaking studies must be carried out to determine the stability of the drug in its final dosage form; that is, its stability to heat, light, and humidity over an extended period of time. Figure 25.7 shows a step in the production of multiple-dose vials under sterile conditions.

Fig. 25.6. Filling of capsules. (*Courtesy of Pfizer, Inc.*).

Fig. 25.7. Production of injectable dosage forms under sterile conditions. (*Courtesy of Pfizer, Inc.*).

Transdermal Drug Delivery Systems

This new dosage form recently was introduced to the pharmaceutical industry. Transdermal preparations now are commercially available for the alleviation of the symptoms of motion sickness (scopolamine), to provide symptomatic relief of angina pectoris (nitroglycerin), and for replacement estrogen therapy (estradiol). Transdermal systems offer advantages commonly associated with continuous intravenous infusion (the bypassing of hepatic "first-pass" metabolism and the maintenance of a constant and prolonged drug level in the body) without the usually attendant risks.

Transderm-Nitro (Fig. 25.8), marketed by CIBA-GEIGY, has a drug reservoir that contains a gel or ointment sealed into the product using liquid filling and sealing technology. The product contains a "rate-controlling" membrane designed to cause the drug to be administered at a predictable, precise rate, thus minimizing intra- and interperson variability.[6]

Advanced Drug Delivery Systems

The industrial pharmaceutical scientist today is not just a formulator of dosage forms but also is a designer of new drug delivery systems.[7] One example is the osmotically controlled release system, which is typically a tablet formed from an osmotically active drug or formulation coated with a semipermeable membrane. One or more minute holes are drilled through the membrane by a laser or high-speed mechanical drill. When the system is exposed to water or any body fluid, water will flow into the core because of an osmotic pressure difference across the membrane. The drug is pumped out of the system through the orifice at a controlled rate. These systems permit prolonged, steady drug release. Additional examples are: the intrauterine device (IUD), which may be employed for the continuous release of very low levels of progesterone; implanted drug pellets, which

Fig. **25.8**. Schematic diagram of arrangement for transdermal medication.

also are used to provide prolonged contraceptive effects; and implanted reservoir and pump systems for the administration of morphine. This is an active area of research, and additional examples may be anticipated.

These new drug delivery systems will permit the consideration of drug candidates that would not have been developed with yesterday's technology and knowledge of drug absorption, distribution, metabolism, and excretion.

QUALITY ASSURANCE—QUALITY CONTROL

In the industrial manufacture of drugs, the variety and complexity of operations make it necessary to assign to a separate and independent group of scientists within the company the responsibility for controlling the quality of the final product. This function is referred to as "Quality Control."

To ensure the required high standards of safety, purity, and effectiveness of drug products, it is necessary to carefully control each step in the manufacturing procedure; this requires control of all raw materials including packaging components and labels, control of the product during manufacture by means of in-process analysis, and to control during packaging of the final product.

The quality assurance and control group also is responsible for ensuring that all products are manufactured and processed in accordance with the current Good Manufacturing Practices (GMP) issued by the FDA, and also with company policies.

As an example, tablets and capsules are examined for the following properties:

- Identity.
- Variability in content of active drug.
- Size.
- Physical appearance.
- Distintegration and/or dissolution time.
- Friability (mechanical stability of tablets).
- Variability in weight.

To determine the variability in the content of the tablet or capsule, a representative sample is taken from each lot, and each tablet or capsule is individually assayed for drug content (Fig. 25.9).

All injectable solutions are tested for sterility and the absence of pyrogens.

Spectroscopic, chromatographic, and other analytical methods are used to determine the purity and the quantity of the drug present in the various dosage forms.

Specifications for containers, closures, and other component parts of drug packages must be established to ensure that they are suitable

Fig. 25.9. Laboratory setup for automatic analysis of individual tablets. (*Courtesy Ciba-Geigy.*)

for their intended use. Also, packaging and labeling operations must be adequately controlled to ensure that correct labeling is employed for the drug. Finished products are identified with lot or control numbers that permit determination of the history of the manufacture and control of the batch.

CONCLUSION

The involvement of the pharmaceutical industry, from basic research to consumer education, constitutes a major contribution to health progress. Largely because of advances in drug therapy, four of the leading causes of death in 1900 no longer prevail. In addition, new drugs have been partially or completely successful in combating many once feared diseases, accounting in good measure for a dramatic decline in death rates in the past three decades. The lower death rates and a decrease in the amount of time lost from work have contributed significantly to our economic well-being. Given the past record, it is reasonable to expect that the pharmaceutical industry will continue to develop and produce new useful products for the betterment of the human condition.

REFERENCES

1. Health Care Financing Administration, Office of the Actuary, "National Health Expenditures, 1988," Health Care Financing Review, **11** (4), 1–41 (Summer 1990).
2. Dimasi, J. A., Hansen, R. W., Grabowsky, H. G., and Lasagna, L., *Journal of Health Economics*, **10**, 107–142 (July 1991).
3. *SCRIP*, Review Issue, 21 (1990).
4. Gilman, A. G., Rall, T. W., Nies, A. S., and Taylor, P. (Eds.), *Goodman and Gilman's The Pharmacological Basis of Therapeutics*, 8th ed., Pergamon Press, Elmsford, NY, 1990.
5. Gennaro, A. R. (Ed.), *Remington's Pharmaceutical Sciences*, 18th ed., Mack Publishing Co., Easton, PA, 1990.
6. Chien, Y. W. (Ed.), *Transdermal Controlled Systemic Medications*, Marcel Dekker, Inc., New York, 1987.
7. Robinson, J. R., and Lee, V. H. L. (Eds.), *Controlled Drug Delivery, Fundamentals and Applications*, 2nd ed., Marcel Dekker, Inc., New York, 1987.
8. DMA 1989-91 Annual Survey Report, 16.

26

Soap, Fatty Acids, and Synthetic Detergents

Shyam Gupta* and Deborah Wiese**

SOAP

Introduction

The origin of the word "soap" is traced to sacrificial Mount Sapo of ancient Roman legend. The mixture of fat and wood ashes that reacted to form soap was carried by rain to the banks of the Tiber river and found as a clay deposit useful for cleaning clothes.[1]

The boiling of fats with ashes was recorded as early as 2500 B.C. Commercial soapmaking was a widespread art in the Middle Ages in Europe. The invention of the soda ash process by Leblanc in 1791, and the discovery by Chevreul in 1811 that soap was composed of a mixture of fatty acids paved the way to modern soap manufacturing processes.[1]

Chemistry

The modern definition of soap relates to cleansing agents derived from fats, oils, and other fatty derivatives: the soaps are alkali and ammonium salts of fatty acids containing from 8 to 22 carbon atoms. These fatty acids are generally a mixture of saturated, unsaturated and polyunsaturated moieties:

$CH_3(CH_2)_{10}COOM$: saturated soap
$CH_3(CH_2)_6CH_2CH = CHCH_2(CH_2)_6$-COOM: mono-unsaturated soap
$CH_3(CH_2)_3(CH_2CH = CH)_2CH_2(CH_2)_6$-COOM: di-unsaturated soap
$CH_3(CH_2CH = CH)_3CH_2(CH_2)_6$-COOM: poly-unsaturated soap
$[M = Na, K, (HOCH_2CH_2)_3NH^+]$

In general, two types of chemical reactions are utilized in the manufacture of soap: the saponification of triglycerides (fats and oils) and the neutralization of fatty acids (which themselves are produced from the triglycerides by a variety of methods, most notably by the splitting of fats and oils with steam). Sodium hydroxide, potassium hydroxide, sodium carbonate, and triethanolamine are the most commonly utilized alkaline moieties in these processes (Fig. 26.1).[2]

In recent years, soap manufacture by an alternate route, the saponification of fatty

*American Soap Company, Olive Branch, MS.
**Monsanto Company, St. Louis, MO.

Fig. 26.1. The chemistry of soap manufacture.

methyl esters, has been under development, most notably in Japan (Lion Corporation) and Italy (Ballestra). The fatty methyl esters are obtained from the methanolysis of triglycerides; inorganic alkali, quaternary ammonium salts, and enzymes (lipase) have been used as catalysts for methanolysis in commercially practiced processes (Fig. 26.2).

Manufacturing Technology

Several designs of commercial equipment, based on the two processes described in Fig. 26.1, are available for the manufacture of soap.

Neutral Fat Saponification	*Fatty Acid Neutralization*
Kettle Processes	Mazzoni SC and SCC
Semi-boiled	Alfa Laval
Boiled	Sato
Modern Processes	Dial Saponiflex
Mazzoni SCNC-N	
Alfa Laval Centipure	
Hitachi	
Dial Saponiflex	

A critical review of these designs is beyond the scope of this chapter. However, the processes themselves are described in a subsequent section.

Raw Materials

The triglycerides (triesters of fatty acids with glycerin) are the basic raw material for the manufacture of soap utilizing a variety of processes. The triglycerides, as a renewable resource, occur widely throughout the plant and the animal kingdoms.

Soap products in the United States are produced almost exclusively from tallow fat and coconut oil or their fatty acid derivatives. The utilization of palm oils, palm kernel oil, and their derivatives for soap manufacture is more commonly practiced in many other parts of the world, especially in Asia. Among other triglycerides cottonseed oil, ricebran oil, castor oil, neem oil, sunflowerseed oil, fish oil, and olive oil also have been used as indigenous feedstocks in many developing nations. The

Fig. 26.2. Fatty methyl ester process for soap manufacture.

fatty acid composition and analytical characterization data of some of the key fats and oils utilized for soap production are given in Table 26.1.

Most of the fats and oils utilized for soap manufacture require a combination of degumming, deodorization, refining, and bleaching steps, which are discussed further in the section on fatty acids.

For both the saponification and the neutralization reactions, sodium hydroxide is the most commonly utilized base. In some special situations, such as the manufacture of transparent and translucent soaps, potassium hydroxide and triethanolamine also are employed.

Functional Properties of Soap

Almost as a rule, the solubility of soap in water increases with an increase in the size of the monovalent cation (base); an increase in the size of a divalent cation (Mg, Ca) results in a decrease in the solubility, however. An increase in the chain length of the fatty acid moiety of soap results in a decrease in the

TABLE 26.1 Fatty Acid Distribution and Analytical Characterization of Soap Base Fatty Acids

Fatty Acid Distribution	*Coconut*	*Stripped Coconut*	*Palm Kernel*	*Palm*	*Tallow*
Caprylic ($C_8H_{16}O_2$)	7	1	4		
Capric ($C_{10}H_{20}O_2$)	6	1	5		
Lauric ($C_{12}H_{24}O_2$)	48	55	50		
Myristic ($C_{14}H_{28}O_2$)	19	22	15	2	2.5
Palmitic ($C_{16}H_{32}O_2$)	9	11	7	42	27
Stearic ($C_{18}H_{36}O_2$)	2	2	2	5	20
Oleic ($C_{18}H_{34}O_2$)	8	7	15	41	42
Linoleic ($C_{18}H_{32}O_2$)	1	1	1	10	5
Linoleic ($C_{18}H_{30}O_2$)					1
C_{20}/C_{22} isomers			1		2.5
Characterization Data:					
Titer, °C	26	29	25	40	40
Iodine value (IV)	10	5	18	55	55
Color, Gardner, max.	5	1			5
Acid value (AV)	270	255			205
Saponification value (SV)	270	255			205
Color, Lovibond, 5-$\frac{1}{4}$"	10Y/1R	4.5Y/ 0.6R			

solubility of soap, whereas the presence of unsaturation results in an increase in its solubility. In general, an increase in the water solubility is accompanied by a increase in the softness and mushiness properties of soap.

For lathering attributes, the use of fatty acids of C_{10} to C_{12} chain length is most preferred. The fatty acids of longer chain length, from C_{16} to C_{18}, contribute to the cleansing properties of soap. A combination of coconut (C_{10}–C_{12}) and tallow (C_{16}–C_{18}) fatty acids thus is most suitable from the viewpoint of consumer-acceptable lathering and cleansing properties of soap.

Manufacturing Processes

The following is a description of the neutral fat saponification and the fatty acid neutralization, two principal processes used for manufacture of soap in the world markets today.

Neutral Fat Saponification. In the saponification of triglycerides with an alkali, the two reactants are rather immiscible. Because saponification is a bimolecular nucleophilic substitution (Sn^2), the kinetic rate enhancement of this reaction is achievable by both a high-shear mixing and an increase in the reaction temperature during the processing. The saponifiction reaction is self-catalyzed to some extent: the formation of soap product effects the emulsification of the two immiscible reactants, causing an acceleration of the reaction rate.

The amount of alkali required to saponify a triglyceride blend is calculated by the following equations:

$$\text{Triglyceride} + 3\text{NaOH} \rightarrow 3\text{RCOONa} + \text{Glycerin}$$

$$\text{NaOH} = [SV \times 0.000713] \times 100/\text{NaOH}(\%)$$

$$[SV/1000] \times [MW(\text{NaOH})/MW(\text{KOH})] \tag{26.1}$$

where SV = saponification value and MW = molecular weight.

The process flow diagram of a typical triglyceride saponification plant is illustrated in Fig. 26.3.[3] The essential components of this plant include proportionating pumps for introducing correct quantities of reaction components into a reactor autoclave, which operates at temperature and pressure conditions appropriate for a rapid and complete reaction. The saponification mixture is recirculated within the autoclave. The temperature of the saponified mass is lowered in a cooling mixer; then the mass is pumped to a static separator for separation of the unwashed soap from the spent lye solution. The soap then is washed with washing lye countercurrently in a washing tower to separate glycerin (as spent lye) from it. Centrifugal separators remove the traces of lye from the soap. The neat soap (60–63% total fatty matter) is neutralized and sent to vacuum spray dryers to produce soap as pellets (78–82% total fatty matter), ready for processing into the finished product on finishing lines.

Soap Drying. The soap mass obtained after the completion of saponification (neat soap) is most commonly dried by a vacuum spray dryer. The water content of soap is thus reduced from 30 to 35 percent in neat soap to 8 to 18 percent in soap pellets or flakes. A variety of vacuum spray dryers, from single-stage to multi-stage designs, are available from several manufacturers. The operation of a single-stage vacuum spray dryer involves the pumping of neat soap through a shell-and-tube heat exchanger where the soap is heated by steam passing through the outside tube in a countercurrent manner. The preheated soap is sprayed onto the walls of a vacuum chamber through a revolving nozzle. The thin layer of dried and partially cooled soap deposited on the walls of the vacuum chamber is removed by rotating scrapers and falls to the bottom of a plodder, which extrudes the soap mass in the form of noodles or pellets.

The multi-stage dryers (Fig. 26.4), which are expanded versions of the single-stage dryers, permit more efficient drying of a wider range of soap formulations than the single-stage dryers.

Fig. 26.3. Continuous neutral fat saponification plant. (*Courtesy G. Mazzoni S.p.A., Busto Arsizio, Italy.*)

Fig. 26.4. Two-stage toilet-soap spray dryer. (*Courtesy G. Mazzoni S.p.A., Busto Arsizio, Italy.*)

Fig. 26.5. Continuous fatty acid neutralization plant. (*Courtesy G. Mazzoni S.p.A., Busto Arsizio, Italy.*)

Fig. 26.6. Toilet soap finishing line. (*Courtesy G. Mazzoni S.p.A., Busto Arsizio, Italy.*)

Fatty Acid Neutralization. The acid–base reaction of fatty acids with alkali to produce soap (below) is much faster than the corresponding reaction of triglycerides with alkalis.

$$RCOOH + NaOH \rightarrow RCOONa + H_2O$$

The amount of alkali (NaOH) required to neutralize a fatty acid blend can be calculated as follows:

$$NaOH = [\text{weight fatty acid} \times 40]/MW\text{fatty acid} \quad (26.3)$$

The average molecular weight of a fatty acid blend is calculated from the following equation:

$$MW\text{fatty acid} = 56.1 \times 1000/AV \quad (26.4)$$

where AV(acid value of fatty acid blend) = mg of KOH required to neutralize 1 g of fatty acid.

A typical Mazzoni fatty acid neutralization soap plant is illustrated in Fig. 26.5.[3] The operation of this plant involves the pumping of the reactants through preheaters to a turbodisperser where the contact of the reactants with each other initiates the formation of neat soap. The soap mass, which is partially reacted at this stage, then proceeds to the mixer where it is recirculated until the neutralization is complete. The completion of neutralization process is monitored by an electric potential (millivolt, mV) measurement for alkalinity. The neat soap then is dried by vacuum spray dryers, as described earlier, to produce soap pellets ready for finishing into soap bars.

Soap Finishing. For the manufacture of bar soap products, the soap pellets are mixed with color, perfume, and other additives in a mixer (amalgamator). The mixed soap mass then proceeds to a series of plodders/roll mills to refine the entire mixture into a uniform, homogeneous product. The finished product is extruded from the last plodder as a continuous slug. A chain cutter with blades cuts it into individual pieces that are stamped with a press into soap bars of the desired size and shape.

The wrapping, cartoning, and end packaging machinery complete the bar soap finishing lines. The schematic diagrams of a variety of soap finishing configurations are illustrated in Fig. 26.6.[3]

FATTY ACIDS

The fatty acids are produced industrially from the triglycerides; this processing involves the steps shown in Fig. 26.7.[4]

Refining involves the purification of triglyceride to remove impurities (phosphatides, polyethylene, chlorophyll, heavy metals, off-odors, off-color) by a combination of acid/alkali washing, clay/activated silica bleaching, deodorization, and hydrogenation steps.

Splitting of triglycerides usually is performed by high pressure steam, resulting in the formation of split fatty acids and glycerin. The production of fatty acids by special "splitting" processes, such as hydrolysis of fatty methyl esters, ozonolysis of unsaturated fatty acids, and chemical oxidation, is practiced in special situations.

Distillation of split fatty acids results in the improvement of chemical, color, and odor quality of the finished product.

Fractionation and Physical Separation

Distilled fatty acid streams are separated to obtain individual components of those fatty acid streams in a higher state of chemical purity. Fractionation is performed to separate the more volatile lower molecular weight fatty acids (C_8–C_{14} chain length); physical separation is practiced to obtain higher molecular weight, less volatile or thermally unstable fatty acids (C_{16}–C_{18} chain length).

Hydrogenation is utilized to reduce the unsaturation and poly-unsaturation of fatty acids. Hydrogenated fatty acids also can be produced by the hydrogenation of triglycerides followed by splitting of the resulting fat/oil blend. The selective reduction of di- and poly-unsaturation often is accomplished by the use of nickel-based specialized catalysts, by a process often called touch hydrogenation.

Fig. 26.7. Production of fatty acids from triglycerides.

In the partial hydrogenation of the unsaturated moieties the isomerization of the double bonds of mono-unsaturated components is a common occurrence; it results in the transformation of the *cis* to the corresponding *trans* stereoisomer. As the physical properties of these isomers may vary significantly, their impact on the properties of the final product should be evaluated rather carefully, for little information is available on this subject at present.

Fat Refining

A typical fat refining plant (Alfa-Laval process) for the acid washing of fats and oils, is illustrated in Fig. 26.8.[5] The triglyceride is degummed to remove phosphatides by acid washing, and then is sent to a bleaching plant where it is dehydrated and treated with an activated clay for the removal of heavy metals, chlorophyll, polyethylene, and color (Fig. 26.9).[3]

Fat Splitting Processes

The direct reaction of triglycerides with alkali followed by acidification or the treatment of triglyceride with water (with or without an acid catalyst) results in the splitting of fats into fatty acids and glycerin. The splitting of fats with water also can be performed by using various lipase enzymes as catalysts. The enzyme process is still in the developmental stages. The inter-esterification of triglycerides with methanol followed by acid or alkali cleavage of the resulting fatty methyl ester is used for the production of certain lower

Fig. 26.8. Acid degumming plant for acid washing of fats and oils. (*Courtesy Alfa-Laval, Tumba, Sweden.*)

Fig. 26.9. Continuous fatty matter bleaching plant. (*Courtesy G. Mazzoni S.p.A., Busto Arsizio, Italy.*)

Fig. 26.10. Fat splitting processes.

molecular weight (C_8–C_{10}) fatty acids; those methyl esters are generally the by-products of fatty alcohol production (Fig. 26.10).

The splitting of fats can be done either in a batch (Twitchell process) or in a continuous (Colgate-emery process) mode; the enzyme-based processes are still under development for commercial applications. A comparison of these processes follows.[7]

Process	*Hydrolysis Time, hr*	*Conversion, %*
Twitchell batch	8–24	95
Continuous	1–3	98–99
Enzyme	72	98

The manufacture of fatty acids from triglycerides via their reaction with alkali followed by acidification is not used commercially. However, the recovery of fatty acids from alkali refining of fats (to produce acid soaps) and soap reboiling operations (the recovery of fatty acids from the acidification of soap waste streams) still are practiced in special situations.

Industrially, the fatty acids are produced predominantly by the high pressure, high temperature splitting of triglycerides. Both batch and continuous processes are available, the continuous process being in greater worldwide use today.

Batch Process. The triglyceride is heated with water at 150 to 250°C and 10 to 25 bar pressure. The splitting is achieved in from 6 to 10 hours, depending on the nature of the triglyceride feed: the higher the molecular weight of the triglyceride, the slower the splitting reaction. The reaction can be

accelerated by increasing the temperature; under high pressure and at 225 to 260°C, no catalyst is required.

A typical commercial batch fat splitting plant (Lurgi) operates at 230°C/400 psig with over 96 percent split yields achieved after two changes of water to remove the glycerin by-product. The sweetwater (glycerin water mixture) is circulated continuously from the base to the top of the autoclave in this process.[6]

A number of other batch fat splitting plants also are available.

Continuous Fat Splitting Process. Continuous fat splitting currently is the most widely used process in the industry. In the single-stage splitting process, fat and water are fed countercurrently into a continuous flow column at up to 260°C/60 bar pressure conditions. The glycerin by-product (sweetwater) produced in the splitting process is removed by water, thus accelerating the reaction equilibrium toward completion (Fig. 26.11).[6] The high splitting temperature ensures adequate dissolution of the water phase in the fat to the extent that mechanical forces for contacting the two reactants are not required. The fat phase passes from bottom to top through the tower, the void volume of the tower acting as a reaction chamber. The water phase travels downward through the mixture of fat and fatty acid. The top and bottom sections of the tower also are designed to effect the transfer of heat from the products to the incoming reactant streams. Split yields of up to 99 percent can be achieved by the use of pre-purified fats.

In summary, the continuous high pressure splitting of fats involves:[7]

- The use of excess water to drive the reaction to completion and for removal of glyerin from the reaction zone.
- The use of a proper combination of temperature and pressure conditions to increase the solubility of water in the fat phase while maintaining the boundaries of the fat/water interface. These conditions also result in an increase in the rate of splitting.
- The possible use of an acid catalyst to accelerate the reaction rate.

Fig. 26.11. Single-stage countercurrent fat splitting plant. (*Courtesy Lurgi GmbH, Frankfurt am Main, Germany.*)

Saponification/Acidification Process. The recovery of fatty acids from soapstocks and soap reboil operations is of commercial value:

$$\underset{\text{(unrefined)}}{\text{Triglyceride}} \xrightarrow[\text{refining}]{\text{caustic}}$$

$$\underset{\text{(refined)}}{\text{Triglyceride}} + \underset{\text{(soapstock)}}{\text{RCOONa}} \xrightarrow{(H_2SO_4)} \text{RCOOH}$$

$$\underset{\text{(impure soap)}}{\text{RCOONa}} \xrightarrow{\text{reboiling}}$$

$$\underset{\text{(purified neat soap)}}{\text{RCOONa}} \xrightarrow{(H_2SO_4)} \underset{\text{(purified fatty acid)}}{\text{RCOOH}}$$

The recovery of fatty acids from soapstocks by a continuous process has been described: soapstocks obtained from degumming and alakli refining operations is subjected to a saponification step followed by controlled acidification for cost efficiency and pollution control.[8]

The saponification/acidification route is used for the manufacture of fatty acids that are sensitive to excessive heat. The splitting of castor oil with NaOH at 100 to 220°C followed by acidification has been used to produce castor fatty acids.[9]

Enzymatic Processes. The enzyme hydrolysis of triglycerides is catalyzed by lipases:

$$\text{Triglyceride} + 3H_2O \xrightarrow{\text{enzyme}} 3\text{RCOOH} + \text{Glycerin}$$

Enzyme-based splitting of triglycerides is under active development and offers the following advantages over modern steam hydrolysis plants:

- The hydrolysis occurs under mild conditions (35–37°C, ambient pressures).
- The color quality of the products is superior.
- The capital investment may be lower.
- The split yields are equivalent.

The selection of a suitable lipase for a commercial viable process is based on the following criteria:[10]

- It should be nonspecific and nonselective as to the structure of the triglyceride.
- It should be stable at above ambient processing temperatures, typically 45 to 65°C.
- It should be economically viable. (This aspect has initiated a number of ongoing studies on the recycling and immobilization aspects of lipase utilization for triglyceride hydrolysis.)[11]

Methyl Ester-Based Processes. The fatty methyl esters are produced predominantly by the transesterification of fats and oils with methanol in the presence of an alkaline catalyst under very mild reaction conditions.[12] They are used in the production of lauric-type (C_{12}) alcohols. The short-chain fatty methyl esters (C_8–C_{10}), produced as by-products via the fractional distillation of crude lauric-type (coconut, palm kernel) methyl esters, are converted to fatty acids via acidic or alkaline hydrolysis (Fig. 26.12). The hydrolysis of short-chain fatty methyl esters by steam splitting or Twitchell-type processes is not very efficient because of unfavorable equilibrium constants.[13]

Fatty Acid Distillation

The removal of contaminants (unreacted triglycerides, odor and color bodies, polymerized matter, and decomposition products) from split fatty acids is achieved by simple distillation.[14] The products obtained from such straight distillations are called "whole cut" fatty acids (whole coconut fatty acid, for example).

The separation of individual fatty acid components from whole fatty acid distillates is achieved by fractional distillation.[15] However, fractional distillation cannot efficiently separate C_{16} to C_{18} saturated from the corresponding unsaturated fatty acids. In such cases, the melting points vary considerably with the degree of unsaturation; and this

Fig. 26.12. Fatty acid production via methyl esters.

property is used for their separation via a number of industrial processes.[16]

Distillation. Straight distillation is used for the removal of low-boiling (top cuts) and high-boiling (pitch) impurities from the split fatty acid stock. The distillation can be done either in a batch or a continuous manner. The batch distillation generally is employed for small-capacity manufacturing; continuous distillation is utilized for large-scale manufacturing and is highly automated (Fig. 26.13).[6]

The split fatty acid (deaerated and preheated) is sent to a distillation column (Fig. 26.13), where it is vaporized rapidly by flowing over heated trays under vacuum (10 mbar) with the addition of live steam (for fatty acid circulation and reduction of the partial pressure). The high boilers are removed by two stages of entrapment: the high boilers and bottoms from the distillation column pass through additional heated trays and a reboiler section to vaporize residual fatty acids from the bottoms; the light end condenser traps the low boilers and the product. The heavy "bottoms" distillate is collected at the bottom portion of the still.

The distillation process is not used for the narrow-cut fractionation. However, by passing a second time through the distillation tower the main-cut distillates may be separated into a light and a heavy fraction; the light end is collected at the top products sections, whereas the heavy end is collected at the bottom residue section of the still. The separation of light and heavy fractions in the first pass is possible for high lauric-type split fatty acids (coconut and palm kernel fatty acids, for example).

The operation of a Lurgi fatty acid distillation system[6] proceeds as follows. The crude fatty acid is dried and degassed in a vacuum dryer/degasser. It then is pumped into the distillation chamber (vacuum of 10–12 mbar) that is divided into 5 to 10 chambers by vertical plates. Each chamber is equipped with a heating element and a baffle. Steam is passed into the system to effect recycling of the liquid between the heating elements and the baffles by the principle of an air-lift pump. The fatty acid is sprayed on the baffles as a fine mist to allow it to evaporate easily; nonvaporized liquid is recycled from the bottom. The fatty acid passes through each chamber in succession, the nondistillable residue being collected in the last chamber, from which it is discharged continuously. The fatty acid vapor is condensed in surface condensers. The typical yields of head cut, refined fraction, and residue are 1 to 5, 85 to 91, and 7 to 11 percent, respectively.

Fractionation. The separation of individual components of a fatty acid mixture into high purity fraction (98%+) is achieved by fractional distillation. In a continuous fractionation still, the fractionation column can be equipped with trays, in the "Thormann"

Fig. 26.13. Continuous fatty acid distillation plant. (*Courtesy Lurgi GmbH, Frankfurt am Main, Germany.*)

system, or with traditional packing to effect contact between the upward-moving vapors and the downward-flowing condensing liquid states.

A typical high-efficiency fractionation system (Fig. 26.14) consists of one continuous straight distillation and one or, preferably, two fractionation columns.[6,14] This system is very efficient for the separation of low-boiling fatty acid components;[3,6] a C_{12} fraction of 99.5 percent purity from coconut fatty acid or a C_{22} fraction of 95 percent purity from rapeseed fatty acid is readily obtained.

From an operational viewpoint, in rectification column 1 of Fig. 26.14, crude fatty acid is dried and vaporized with the addition of

Fig. 26.14. Fatty acid fractionation. (*Courtesy Lurgi GmbH, Frankfurt am Main, Germany.*)

steam under a vacuum of 8 to 40 mbar. A mixed fraction is withdrawn from the column top. The product from the bottom of column 1 is vaporized in rectification column 2 with steam under a pressure of 6 to 8 mbar. The low-boiling fraction is collected in downstream scrubbers after the vapor has been condensed in surface condensers.

The separation scheme for coconut fatty acid is shown in Fig. 26.7. A variety of alternate fractionation schemes can readily be devised.

Fatty acid distillation has been done for over 100 years.[14] However, most of the currently practiced art has originated in the past 25 years.[15] The majority of the recent work falls in the following areas:

- Improvements in the column design to reduce residence time.
- Enhancements of system design to minimize pressure drops.
- Demister designs for better air/effluent water pollution control.

Special Separation Methods

The separation of long-chain saturated fatty acids (C_{16}–C_{18}) from the corresponding unsaturated fatty acids by distillation is not practical because of the proximity of their boiling points. However, both the melting points and the solubility (in organic solvents) of these two types of fatty acids are vastly different.[17] These properties are used to advantage for separating mixtures of saturated and unsaturated fatty acids into fractions that are rich in either saturated (stearin) or unsaturated (olein) components.[16]

$$\text{Tallow fatty acid} \xrightarrow{\text{separation}} \underset{\text{(Stearin)}}{\text{Stearic acid}} + \underset{\text{(Olein)}}{\text{Oleic acid}}$$

The following processes are used commercially for the separation of such fatty acids: mechanical pressing, solvent crystallization, and hydrophilization.

Mechanical Pressing. This method, also called panning and pressing, is the simplest and oldest method, and is still practiced in many Third World countries. In the method, fatty acids is melted in shallow pans and then allowed to cool. The material transforms into a crystalline solid, with a liquid phase trapped between the crystals of the solid phase. The

liquid portion (olein) is squeezed out by the application of mechanical force, separating it from the solid (stearin) component.

The pressing can be done by hydraulic presses in either a batch or a continuous mode. A multiple number of melting/crystallization/pressing cycles is utilized for obtaining fractions of increasing purity:

$$\text{Tallow fatty acid} \xrightarrow{\text{first pass}} \underset{\text{(Stearic acid)}}{\text{Single press}} \xrightarrow{\text{second pass}}$$

$$\underset{\text{(Stearic acid)}}{\text{Double press}} \xrightarrow{\text{third pass}} \underset{\text{(Stearic acid)}}{\text{Triple press}}$$

The unsaturation, as determined by the iodine value (IV), decreases from an IV of 10.0 for a single press to an IV of 1.0 for a triple press stearic acid grade.

Solvent Crystallization. Two processes, one utilizing acetone (Armour) and the other methanol (Emery), are well known. Using the latter, fatty acid is dissolved in 90 percent aqueous methanol in a 1:2 acid/solvent ratio by the application of heat. The resulting solution then is cooled to −15°C in a multi-tubular crystallization chamber equipped with scrapers for efficient heat transfer. The crystallized fatty acids are removed by filtration. The filter cake is melted and stripped of any residual solvent to give the refined stearin fraction, and then the liquid stearin is converted to flakes or powder by a variety of processes. The mother liquor from the above filtration is stripped to obtain the olein fraction. The separated stearin and olein fractions have a variety of commercial applications in both chemical and food industries.

Hydrophilization Process. In this process, olein and stearin are separated by forming a suspension of the stearin moiety in water with the aid of a surfactant followed by filtration to remove the stearin (Fig. 26.15). The mixture of fatty acids is melted and then partially crystallized by cooling. The resulting slurry, which is a mixture of olein (liquid) and stearin (crystalline solid), is mixed with water containing a wetting agent (sodium dodecylbenzene sulfonate, sodium lauryl sulfate). After mixing for a specified period, the slurry is centrifuged; the light fraction provides the liquid olein moiety, the middle fraction is a suspension of solid stearin in water, and the bottom portion is water/surfactant solution.[6]

Thermodynamically, the fatty acid crystals, after being wetted by the surfactant solution, transfer from the liquid oil into water solution. The oil droplets then coalesce to give the lighter olein phase. The separated stearin is remelted to remove residual water/surfactant, and then reprocessed to the desired level of purity.

Although the quality of the final product from these three processes is fairly comparable, hydrophilization currently is the most widely

Fig. 26.15. Hydrophilization process for separation of olein and stearin. (*Courtesy Lurgi GmbH, Frankfurt am Main, Germany.*)

used process for the separation of stearic and oleic acids.

SYNTHETIC DETERGENTS

Introduction

Synthetic detergents initially were developed as soap substitutes in an economy that was running short of edible fats and oils. These synthesized compounds were more resistant to calcium salt formation in hard water and generally provided improved wetting, cleansing, and surfactancy over soaps. The term "synthetic detergents" has been shortened to "syndets" to describe detergent compositions that are comprised of the synthetic active ingredient, water hardness control agents (builders), and other detergent additives. "Surface active agent" has been shortened to "surfactant" to describe the surface active principal or active ingredient (AI). Since their introduction in the 1940s, synthetic detergents have continued to grow in both total volume and percentage of detergent use (Table 26.2).

Whether in liquid or in powdered forms, present detergent products are complex mixtures of several ingredients. The formulations consist primarily of surfactants, builders, and other additives designed to maximize performance for the consumer while maintaining reasonable raw material and manufacturing costs. Recent trends in both liquid and powdered product forms show increased utilization of mixed builder and surfactant systems.[18,19] Also, performance additives such as bleaches, bleach activators, softeners, and enzymes often are present in the detergent to boost overall performance. However, in all products, the backbone of the detergent is the surfactant/builder system.

TABLE 26.2 U.S. Consumption of Soap and Synthetic Detergent[a] (millions of pounds)

Year	*Soap*	*Synthetic Detergent*	*Total*
1940	2306	30	3236
1950	2882	1443	4325
1960	1230	3940	5170
1970	1050	5650	6700
1980*	1300	6400	7700
1990*	1625	9000	10,625

[a]CEH Marketing Report, "Speciality Chemicals," *Chemical Economics Handbook*, SRI International, Aug. 1988.
*Estimated.

Surfactant Hydrophile-Hydrophobe Balance

As the term indicates, surface active agents are active at surfaces by preferential orientation of the molecule. This suggests that some built-in characteristic of the surfactant controls or contributes to molecular activity. Surfactants can be tailored for effective use in either aqueous or nonaqueous systems, depending upon their solubility characteristics. An optimum hydrophile–hydrophobe balance generally exists for a specific detergent application and for the compound class used. This optimum composition often is developed through evaluation of required hydrophil–hydrophobe characteristics of the application. For example, alkylbenzene is slightly surface active in nonaqueous media but insoluble and ineffective in water. Addition of a single SO_3^- group through alkylbenzene sulfonation provides high water solubility and excellent surfactant characteristics in water, but insolubility in petroleum solvents. If alkylbenzene is di- or trisulfonated, the compound becomes more water-soluble and loses much of its surface activity, operating essentially as an additional electrolyte in the system. In both cases, the hydrophile balance is exceeded, with insufficient hydrophobe influence to increase the preferential orientation necessary for surfactancy. However, neutralization of the SO_3H group of highly water-soluble alkyl benzene sulfonic acid with a long chain amine produces an effective nonaqueous surfactant. A similar effect of decreasing water solubility and increasing nonaqueous surfactancy can be achieved by increasing the alkylbenzene carbon side chain length from C_{12} to C_{18} or higher.

Nonionic surfactant systems depend upon a multiplicity of oxygen and hydrogen bond linkages to effect hydrophile–hydrophobe balances similar to those exhibited by SO_3

solubilizing groups in alkylbenzene sulfonate. Ethylene oxide–adduct types of nonionics can be varied in chain length to impact hydrophobe characteristics. Additionally, adjustments in degree of ethoxylation can be used to modify the hydrophile balance.

In studying the literature, one might believe that lauryl sulfate, sodium stearate, dodecylbenzene sulfonate, and so on, are pure compounds. Commercially available compounds are mixtures of homologs, with lauryl sulfate being a generic designation for a mixture of sulfates whose largest fraction is derived from the C_{12} alcohol; the remainder come from higher and lower alcohols. The total distribution of homologs is dependent upon the degree of fractionation in the original alcohol production. The same is true for alkylbenzene derivatives and nonylphenol compounds, while soaps are mixtures of the various fatty acids natural to fats and oils. Nonionic ethylene oxide adducts are further examples of mixtures; the nonaethylene glycol ether designation suggests that this compound is the primary constituent, but ethylene oxide adds to individual hydrophobe molecules in a manner that can yield a normal Poisson distribution of adducts, with the 9-molar adduct representing the distribution mean. Although it might seem that a competitive edge could be gained by supplying a highly purified product, mixtures frequently permit usage where pure compounds are less effective. Long experience with soaps and synthetics shows that except for highly specialized purposes, pure compounds are not competitive with properly chosen soap or synthetic mixtures.

Anionic Surfactants

Alkylaryl Sulfonates. Alkylbenzene sulfonates have been the "workhorse" of the detergent industry for more than 25 years. They account for approximately 50 percent of the total synthetic anionic detergent volume used in liquid and spray-dried formulations. Three basic grades of alkylate (alkylbenzene) are manufactured, with molecular weight ranges of approximately 235, 240, and 260. In general, the C_{10-12} range alkylates are used for light duty liquid systems. The C_{12-14} range alkylates are used in heavy duty liquid and spray-dried detergents. The alkylate is sulfonated and neutralized, primarily to a sodium salt, prior to use in detergent formulations.

Prior to 1965, alkylbenzene production was synthesized from petroleum tetrapropylene reacted with an aluminium chloride or hydrogen fluoride catalyst and benzene. The resultant alkylate was a "hard" branched-chain compound that was considered slowly biodegradable. A straight-chain alkylate, termed LAB (linear alkylbenzene), has been produced since 1965 in the United States. Extensive research has demonstrated biodegradation effectiveness in sewage treatment plants in excess of 95 percent.[20–23]

Three basic processes have been practiced for linear alkylbenzene manufacture. The most prevalent route of alkylbenzene manufacture is by partial dehydrogenation of paraffins, followed by alkylation of benzene with a mixed olefin/paraffin feedstock, using liquid hydrogen fluoride catalyst. A second route is via partial chlorination of paraffins, followed by alkylation of the chloroparaffin/paraffin feedstock in the presence of an aluminium chloride catalyst. The third process uses partial chlorination, but includes a dehydrochlorination to olefin step prior to alkylation with aluminium chloride or hydrogen fluoride.

Distillation is used in all processes to separate alkylation reaction products into the targeted detergent alkylate homolog distributions. A coproduct stream of high molecular weight linear alkylbenzene and dialkylbenzenes also is produced, for use in lubricant detergents and industrial applications.

Products with low 2-phenyl isomer content (13–22%) are produced by hydrogen fluoride alkylation of internal olefins from paraffin dehydrogenation. High 2-phenyl products (25–35%) are produced from aluminium chloride alkylation of chloroparaffins. Generally, high 2-phenyl products are preferred for use in liquid detergent applications because of the improved solubility found with the higher 2-phenyl level. Low 2-phenyl alkylates

also can be used for this application, with additional usage of hydrotropes.

Sulfonation. For detergent use, the alkylate must be sulfonated to an acid form and then neutralized with a base such as sodium hydroxide. The major uses of the neutralized slurry are in the production of spray-dried detergent and both light and heavy duty liquid detergents. The slurry also can be drum-dried to a powder or flake, or spray-dried to light-density granules. Dried forms are used primarily for institutional and industrial cleaning applications and car washing compounds. The sulfonation reaction is quite exothermic and almost instantaneous. In order to prevent decomposition and maintain optimum color of the resultant sulfonic acid, an efficient reactor heat removal system is necessary.

Oleum and sulfur trioxide are used routinely for sulfonation, with the latter growing in popularity in recent years. Oleum sulfonation requires relatively inexpensive equipment and can be accomplished by either batch or continuous processes. However, the oleum process has major disadvantages in the SO_3 cost/lb (as H_2SO_4), the need for spent acid stream disposal, and potential corrosion problems due to sulfuric acid generation. The oleum process typically yields an 88 to 91 percent purity sulfonic acid with the remainder consisting of approximately 6 to 10 percent H_2SO_4, 0.5 to 1.5 percent water, and 0.5 to 1.0 percent unsulfonated oils. Reactions involved in oleum sulfonation are shown below.

Sulfonation:

Alkylbenzene + Oleum (20–25%)

Sulfonic acid + 96% Sulfuric acid

Neutralization:

Sulfonic acid + Sulfuric acid + Caustic

Sodium alkylbenzene sulfonate + Sodium sulfate + Water

A typical batch oleum sulfonation process is shown in Figure 26.16. Four basic processing operations are utilized prior to neutralization: sulfonation, digestion, dilution, and phase separation. Mixing of the alkylate and oleum is done in the sulfonation stage. Heat removal from the exothermic reaction is achieved by reactor jacketing and/or recirculating heat exchanger loops. Key variables in controlling alkylbenzene conversion and acid color are the temperature, acid strength, reaction time, and oleum-to-alkylate ratio. The product leaving the sulfonation zone is aged or digested 15 to 30 minutes to allow reaction completion; at this point, the reaction product is a mixture of sulfonic acid and sulfuric acid. Following digestion, the mixture is diluted with water to quench the reaction. The reaction mass then is fed to a separator to allow gravity settling of the spent sulfuric acid layer from the lighter sulfonic acid. The lower spent acid layer contains approximately 75 to 80 percent sulfuric acid; and the upper layer consists of approximately 88 to 91 percent sulfonic acid, 0.5 to 1 percent unsulfonated oils, and 6 to 10 percent sulfuric acid. The separation time is dependent upon the separator vessel configuration, sulfonic acid viscosity, temperature, and degree of aeration in feed streams.

The gaseous air/SO_3 sulfonation process normally generates a product composed of 95 to 98 percent purity sulfonic acid with 1 to 2 percent sulfuric acid, 0.5 to 1.5 percent water, and 1.0 to 2.5 percent unsulfonated oils.

Fig. 26.16. Batch oleum sulfonation process.

Although the initial capital investment costs are much higher than those for an oleum process, relative sulfonic acid yields and spent acid disposal costs are substantially lower. The chemical reactions involved in air/SO_3 sulfonation are as follows:

Sulfonation:

Alkylbenzene + Sulfur trioxide — Alkylbenzene sulfonic acid — + Anhydrides

Hydration:

Anhydride + Water — Alkylbenzene sulfonic acid

Neutralization:

Alkylbenzene sulfonic acid + Sulfuric acid + Caustic — Sodium alkylbenzene sulfonate + Sodium sulfate + Water

Fig. 26.17. Alkyl benzene sulfonation schematic flow diagram.

A typical continuous air/SO_3 sulfonation process is outlined in Fig. 26.17. Sulfur trioxide for the sulfonation reaction is either produced from vaporization of liquid SO_3 or generated on-site by burning sulfur and then catalytically converting SO_2 to SO_3 over a vanadium pentoxide catalyst. Increased shipping regulations and sensitivity to storage of the highly hazardous liquid SO_3 have increased the trend toward on-site sulfur trioxide generation. The vaporized sulfur trioxide is diluted to 4 to 8 wt. % in dry air prior to sulfonation. The air serves as a diluent and heat sink for the highly exothermic sulfonation reaction.

In the sulfonation reactor, the molar ratio of SO_3/alkylate is controlled between 1.03/1.0 and 1.06/1.0 to achieve optimum conversion to sulfonic acid without increased side reactions or color degradation. The Chemithon,[24,25] Ballestra,[26] and Mazzoni[27] units have unique falling film reactor designs to provide intimate mixing of the alkylate and air/SO_3 streams. The reactor temperature is controlled between 110 and 150°F, depending upon the desired production rates and product quality targets. Digestion and hydrolysis temperatures are controlled between 110 and 125°F. During hydrolysis, approximately 2 parts of water per 100 parts of reaction product are added to convert acid anhydrides to sulfonic and sulfuric acid.

Sulfur trioxide sulfonation also is practiced in processes[28,29] that use liquid SO_2 as a refrigerant and a solvent. The product quality is high, and the sulfonic acid color is low because of low temperatures and rapid removal of the reaction exotherm via SO_2 vaporization. Equipment consists of fairly simple glass-lined reactors equipped with gaseous SO_2 recovery. Control of SO_2 recovery, frothing, and purging of SO_2 from the reaction products are keys to achieving a high-quality product. As in the case of air/SO_3 sulfonation, this process can be used for all detergent-grade surfactants, including alkylbenzene, alcohol ethoxylates, and alcohols.

A novel venturi reactor has also been patented by Chemithon Corporation.[30,31] The organic reactant is sulfonated by injection into a stream of gas containing sulfur trioxide at a venturi. The reaction mixture is quenched and recycled downstream until sulfonation is complete. This reactor design is utilized primarily for highly viscous, high molecular

weight specialty surfactants for industrial applications such as lubrication oil additives.

Neutralization of sulfonic acid is similar in the case of either an oleum or a SO_3 sulfonation process. The sulfonic acid can be neutralized with aqueous solutions of base; that is, KOH, NH_4OH, NaOH, or alkanolamines. The sodium salts are used for formulating spray-dried detergents for household laundry consumption. Ammonium and alkanolamine neutralized salts generally are used in light duty liquid detergents. In general, neutralized salts from oleum sulfonation contain approximately 8 to 10 percent sodium sulfate, where as those from air/SO_3 sulfonation contain only 2 to 3 percent sodium sulfate.

Sulfation. For sulfation of alcohols or ethoxylated alcohols, falling film reactors also are used, but cooling water and sulfation temperatures are adjusted to lower values for these raw materials. The digestion and hydration steps are not required in sulfation, and these systems are bypassed. The acids are neutralized immediately to minimize degradation and side reactions in storage.

Typical sulfation reactions for alcohols and alcohol ethoxylates are shown in the following equations:

$$R—OH + SO_3 \rightarrow R—O—SO_3H$$

$$R—(OC_2H_4)_n—OH + SO_3 \rightarrow R—(OC_2H_4)_n—O—SO_3H$$

Typical processes for air/SO_3 sulfonation of alpha olefins are shown in Fig. 26.18 (alpha olefins), and for air/SO_3 sulfation of fatty acohols in Fig. 26.19. Bleaching often is required in the post-hydrolysis steps to remove color bodies prior to neutralization.

For laundry detergents containing mixtures of alkylbenzene sulfonate, alcohol sulfates, or alcohol ethoxy sulfates, sulfonation reactions are either campaigned or produced in two separate units—one for sulfonation, one for sulfation. The respective acids can be blended prior to neutralization to yield a superior product to one made by post-blending of the neutral salts.

Neutralization. Sulfonic acid also can be processed to a dry powdered form by

Fig. 26.18. Alpha-olefin sulfonation schematic flow diagram.

Fig. 26.19. Fatty alcohol sulfation schematic flow diagram.

stoichiometric neutralization with a base salt, such as sodium carbonate. The neutralized mixture is slightly moist at this stage, and requires blending with other detergent additives to complete its transition to a dry free-flowing powdered detergent. The Ballestra[32] process mixes the acid and salts together, whereas the Staffer[33] process meters the acid and caustic streams into a tumbling product bed. The A. G. Hoechst[34] process degasses the sulfonic acid prior to neutralization. Removal of SO_2 or SO_3 gas and improved temperature control produce a neutralized paste of excellent color quality.

Nonionic Surfactants

Nonionic surfactants are increasingly popular active ingredients for liquid heavy duty laundry detergent formulations. The majority of nonionic surfactants are polyoxyethylene or polyoxypropylene derivatives of alkylphenol, fatty acids, alcohols, and amides. It has been estimated that approximately one ethylene oxide unit is required to solubilize each methylene unit. Thus, by changing the ratio of ethylene oxide to fatty acid, the nature of the surfactant can be modified from an oil-soluble to a water-soluble product. This versatility makes applications possible in numerous areas, such as detergents, agricultural products, metal protection, emulsion polymerization, textile applications, deinking, and penetrants. The reaction of ethylene oxide with fatty acids was described initially in 1928. Shick[35] and Jungermann[36] have edited very thorough volumes concerning nonionic surfactants.

Ethoxylation. In processing, the fatty acid and catalyst are charged to a reactor and purged with an inert gas such as nitrogen. The partial pressure is adjusted so that the explosive limits of the gas mixture in the reactor during and after addition of ethylene oxide are not exceeded. The mass is heated to reaction temperatures of 115 to 200°C, and ethylene oxide is added until the desired pressure is reached. Heat removal rates control the rate of reaction. Excessive cooling can result in excessive absorption of ethylene oxide into the reaction mixture, which can result in sudden reaction followed by overpressure. The reactor temperature and pressure are controlled until the desired quantity of ethylene oxide has been added. Then the

ethylene oxide addition is discontinued, the batch is cooled, and the catalysts are removed by filtration.

Other examples of nonionics commonly used in syndets are glycols, glycerols, sugar esters, and alkanolamides. Polyethylene glycol esters of fatty acids are produced by reaction with ethylene oxide or by esterification of the fatty acid with polyethylene glycols. The reaction with ethylene oxide is described as:

$$RCOOH + nC_2H_4O \rightarrow RCOO(C_2H_4O)_nH$$

Lauric, palmitic, stearic, and oleic acid also are ethoxylated and widely used as detergents and emulsifiers.

Amphoteric Surfactants

These surfactants contain both cationic and anionic groups. The products are relatively expensive because of the raw materials involved and processing costs. Typical applications include shampoos, bubble baths, and other toiletries and cosmetics where mildness is a major concern. A typical product is shown below:

N-Beta-hydroxyethyl-*N*-beta-carboxymethyl fatty acid amidoethylamine, sodium salt

Other examples of fatty amine oxides include dimethylcarboxymethyl coconut fatty acid, propylamide ammonium betaine, alkylamide betaine, and cocamidopropylamine oxide. Additional varieties are cited by McCutcheon.[37]

Alkylolamides

A wide variety of amine condensates can function as foam boosters, wetting agents in aqueous/nonaqueous systems, emulsifiers, and dispersants. These are mainly reaction products of diethanolamine (DEA) or monoethanolamide (MEA) with fatty acids at a 2:1 ratio. Excess DEA converts both the amino and the amido ester to an active 2:1 product, *N*,*N*-bis(2-hydroxyethyl) lauramide. In detergents, alkylamides are used at 1 to 2 percent levels as foam stabilizers. The general formula for alkylanolamide is:

R^1 = alkyl; R^2 = hydrogen, alkyl, or hydroxyalkyl; and R^{2-3} = hydrogen or alkyl

Cationic Surfactants

Cationic surfactants tend to be specialty items that are tailored for diverse uses: germicides, textile applications, sensitizers, flotation agents corrosion inhibitors, and fabric softeners. Their primary growth has been in fabric softeners for home laundry detergent. Because cationics are antistatic agents, they perform well in removing static charges associated with synthetic fabrics.

These surfactants have at least one hydrophobic group attached directly to a positively charged nitrogen molecule. Cationic surfactants may be formed from nitriles, amines, amide-linked amines, or quaternary nitrogen bases. Significant literature reviews[38,39] cover this subject in detail.

Detergent Builders

Builders typically are added to a detergent formulation to extend or improve the formulation cleaning performance across a wide range of use conditions. The combination of builders and surfactants exhibits a synergistic effect to boost total detergency and cleaning efficacy, as compared with an equal amount of either compound alone. The major properties and characteristics that are desirable for a compound to be considered a detergent builder include:

1. The ability to control water hardness and other metal ions.
2. A contribution to final product alkalinity.
3. Buffer capacity in the proper pH range.
4. Deflocculation capability.

5. Compatibility with other formulation ingredients and detergent additives.
6. Consumer safety.
7. Environmental acceptability.
8. Processibility.
9. Adequate cost/performance.

In today's detergent products, a variety of builder materials are used singly or in combination with other builders to contribute unique properties for final product performance enhancement.

Phosphates. Tetrasodium pyrophosphate ($Na_4P_2O_7$—TSPP) became commercially available in the mid-1930s and soon gained acceptance as the preferred builder for soap-formulated products. The combination of TSPP and sodium carbonate was more effective than either ingredient used separately. By the mid-1940s sodium tripolyphosphate ($Na_5P_3O_{10}$—STP) was commercially available and had supplanted TSPP as the preferred builder because of superior detergent processing, solubility, and hardness ion sequestering characteristics. Sequestration is defined as the reaction of a cation with an anion to form a soluble complex. The sequestration of Ca^{++} and Mg^{++} ions leads to softened water and is the most important function of any detergent builder.

Sodium tripolyphosphate is the major builder ingredient in heavy duty laundry detergents, automatic dishwashing compounds, and industrial and institutional cleaners. In liquid consumer laundry detergents, sodium citrate has displaced tetrapotassium pyrophosphate ($K_4P_2O_7$—TKPP), but TKPP use in industrial and institutional liquids is preferred because of its superior solubility and sequestration characteristics.

Other possible builder salts include sodium trimetaphosphate, trisodium and tripotassium phosphates, and sodium acid pyrophosphate. Each has a unique property whose utility is dependent upon the particular detergent system or application. The book *Phosphorus and Its Compounds*[40] presents a complete assessment of phosphates and their role in detergents.

Silicates. Both sodium and potassium silicate solids or solutions have valued functional characteristics such as emulsification, buffering, deflocculation, and anti-redeposition ability. Silicates also provide corrosion protection to metal parts in washing appliances, as well as to the surfaces of china patterns and metal utensils in automatic dishwashers. Silicates are manufactured in either liquid, crystalline, or powdered form; and they also vary in degree of alkalinity. The alkalinity of the silicate provides buffering capacity in the presence of acid soils and enhances the sequestration ability of the builder system in the formulation. The silicate/alkali ratios of the silicate are selected by the formulator to meet specific product requirements. Silicate ratios of 1/1 commonly are used in dry blending applications, with liquid silicates with silicate/alkali ratios of 2/1 and higher commonly used in laundry and automatic-dishwashing applications.

In using liquid silicates, it is essential that the proper silicate ratio be selected for various process and performance parameters. For example, liquid silicates containing a high SiO_2/Na_2O ratio are susceptible to forming insolubles if overdried. In agglomeration processes, the use of lower SiO_2/Na_2O ratios of liquid silicates requires additional mixing, aging, and drying to obtain a crisp, free-flowing product.

Carbonates. In certain areas of the United States, phosphate sales have been legislatively banned from use in laundry detergent products. In these areas, sodium carbonate (Na_2CO_3) or a combination of sodium carbonate and zeolite has replaced STP as the builder in granular laundry products. Sodium carbonate softens water by precipitating the hardness minerals, and it is one of the most economical sources of alkalinity for detergents and provides some degree of soil dispersion and suspending action. Sodium carbonate is commonly used in powdered laundry detergent, automatic-dishwashing compounds, hard surface cleaners, and presoak formulations.

Sodium Citrate. Sodiumcitrate ($Na_3C_6H_{25}O_7$) has builder characteristics due to its ability to sequester water hardness ions and deflocculate soils. Sodium citrate's principal use is in liquid laundry products, with limited use in dry powders.

Zeolites. An important alternative builder for powdered laundry detergents in regions where phosphate salts have been legislatively banned for laundry product use, zeolites are crytalline hydrated aluminosilicates of Group I and Group II elements. They are sodium, calcium, magnesium, potassium, strontium, and barium salts, and appear in natural or synthesized forms. Zeolite Type 4A (empirical formula—$Na_2O \cdot Al_2O_3 \cdot 2SiO_2 \cdot 4.5H_2O$, commonly called Zeolite A) is the most widely used form for laundry detergents.

Zeolites are not water-soluble, but they reduce calcium water hardness by ion exchange. The calcium ion passes through the zeolite pore openings and exchanges with the sodium ion. Magnesium ions are not significantly removed by zeolites because of the large radius of magnesium; therefore, zeolites are used in conjunction with other builders such as sodium carbonate to control magnesium and other ions present in the wash water.

Sodium Nitrilotriacetate. One of the first compounds developed as a replacement for sodium tripolyphosphate was sodium nitrilitriacetate ($N(CH_2COONa)_3 \cdot H_2O$—NTA), which has excellent sequestration and/or chelating ability. This product was quickly accepted by the detergent industry as an alternate builder, but its use in detergents was suspended in the United States in December 1970, when three major detergent manufacturers voluntarily agreed to discontinue its U.S. usage because of preliminary evidence indicating that the compound could be teratogenic. Subsequent assessment (May 1980) by the EPA, culminating nearly 10 years of extensive safety studies, indicated that the risk to persons exposed to NTA from detergents is not significant. However, the EPA stated that NTA should not be used in products such as shampoos, hand-dishwashing detergents, foods, or other consumer products designed for direct dermal or oral exposure. Although NTA has not been used in U.S. household laundry products since the voluntary ban except in brief test markets, it has been widely used as a detergent builder in Canada since 1970, and in the United States in industrial and institutional cleaning products. NTA normally is used in liquids as the primary builder or as a supplement to phosphates in spray-dried detergents.

Detergent Additives

In addition to detergent builders and surfactants, there are numerous additives that perform other specific functions in detergents. A few of the most widely used materials are described below.

Sodium Sulfate. Although considered a builder by some manufacturers, sodium sulfate (Na_2SO_4) contributes little to detergent performance. It is commercially available from natural sources and as a by-product from rayon processing. In dry mix and agglomerated type products, sodium sulfate is used to improve finished product flow characteristics. In spray-dried products, sodium sulfate acts as an inert "filler" and aids in density control and flow characteristics.

Sodium Chloride. Sodium chloride (NaCl) still is used by some formulators in detergent manufacture, its main function being as an inert "filler" or diluent. In spray-drying applications, NaCl is used to control slurry viscosity and the density of the spray-dried bead or graules. Although salt is used to some extent in automatic dishwashing formulations, its use is not recommended[41] because of potential machine and dishware corrosion.

Sodium Carboxymethylcellulose (NaCMC). The surfactants employed in synthetic detergents are very effective in removing soil and stains from the surface of the washed fabric. Under heavy soil loading, there is a tendency for these soils to redeposit on the fabric.

NaCMC is effective in reducing soil redeposition at formulation levels of 0.1 to 0.5 percent.

Optical Brighteners. Although the surfactant system in detergents cleans clothes effectively, white fabric tends to "gray" or lose its white appearance after numerous washes. Optical brighteners are added to detergent formulations to visually improve the "whiteness" or brightness of the fabric.

All major liquid or powdered detergent products contain one or more fluorescent whitening agents.[42] Substantive to most fabrics, these materials absorb ultraviolet radiation and convert it to a visible blue-white reflectance range. This characteristics masks the natural yellowing or graying of the fabric. Most brightener systems are more effective on cotton than on synthetic fabrics. The usage range varies from 0.05 to 0.2 percent in the detergent formulation.

Hydrotropes. Hydrotropes are used primarily in liquid detergents to increase the solubility of less soluble ingredients and improve storage stability. They also are used occasionally as viscosity modifiers in the slurry phase of manufacturing spray-dried detergent powders. Reduction of the detergent slurry viscosity improves its processing characteristics, reduces product bulk density, and ensures crisp, free-flowing granules. The most common hydrotropes are the ammonium, potassium, or sodium salts of toluene, xylene, or cumene sulfonates.

Enzymes. Proteolytic and amyolytic enzymes have been designed for use as adjuvants in detergent systems to remove stains. These complete proteinaceous molecules act as catalysts and tend to break down particular soils and stains to a form more readily removed from fabrics. These ingredients are used in presoak detergent systems and some liquid and powdered laundry detergents.

Pearlescent–Opacifying–Thickening Agents. These compounds are used in liquid detergent systems to produce a specific aesthetic effect. Pearlescent compounds (bismuth, titanium dioxide–coated mica) normally are insoluble in water or alcohol, but are dispersed in emulsion systems. Opacifying agents (e.g., water-soluble salts of styrene and maleic anhydride) are used to reduce translucence, modify the viscosity characteristics of the liquids, and provide a creamlike texture or pearlescent effect to the product.

Thickeners such as bentonite clays are used in providing rheological modification for automatic-dishwashing liquids. The clays impart a thixotropic nature to the liquids that allows sufficient yield stress to prevent premature dispensing into the automatic dishwasher. The use level in automatic-dishwashing liquids is at 0.5 to 1.0 percent.

Anticaking Agents. Because of the reduced phosphate levels and use of nonphosphate builders in some powdered laundry detergents, flow characteristics and caking tendencies have changed. Products formulated with increased anionic active levels and higher sodium carbonate content tend to exhibit compaction caking in the detergent carton. Anticaking agents have been added to the formula to reduce this tendency, especially in warm humid climates. A wide range of products are available, including sodium benzoate, tricalcium phosphate, colloidal aluminium oxide, silicon dioxide, magnesium silicate, calcium stearate, and microcrystalline cellulose.

Spray Drying

A major portion of the powdered laundry detergents today are spray-dried. In this process, the detergent raw materials are mixed with water to form a thick paste or slurry, atomized into spherical droplets, and dried to produce a crisp, free-flowing graular product. The art of spray drying was first patented in 1883 but was not practiced to any extent until the 1930s. With the advent of synthetic detergents in the 1940s spray drying became the dominant method of detergent processing. The advantages of spray drying are numerous; for example, the product density can be varied from 0.20 to 0.80 g/cc, the detergent granule

solubility is significantly improved, the production rates are increased, and multiple formulations can be prepared using the same equipment. The major disadvantages of spray drying are the initial capital investment required, its relative inflexibility in varying products sizing and density in some units, and the energy-intensive nature of the process. Spray drying is divided into several process steps: selection of formulations, slurry preparation, atomization, drying, conditioning of the product, post-additions, packaging, waste reclamation, and pollution control.

Spray-dried detergents are manufactured for use in locations where various legislative restrictions may limit phosphate use. Typical formulations are shown in Table 26.3.

Slurry preparation of the detergent formulation is considered to be the single most important factor in spray drying. Control of the raw material order of addition, slurry solids level, slurry temperatures, viscosity, and aeration enables manufacturers to deliver the desired final product characteristics. Essentially, the spray tower itself is simply a contained heat source that removes moisture at a uniform rate dependent upon the quality and the quantity of feed to the dryer. If feed variables are not controlled, variability in the density, moisture level, aging, and/or packaging characteristics increases. For example, if the slurry solids level is not constant ($\pm 2\%$), the evaporative load, production rate, particle sizing, and moisture levels of the spray-dried granules will vary. Compensating for nonuniform feed to the spray tower is difficult because of the interaction of controllable variables such as tower temperature, pump pressure, gas velocity, and so on. Although some minor tower adjustments sometimes are necessary to control process conditions, the major control for routine operations is the preparation and uniformity of the detergent slurry.

Slurries are prepared via batch, semicontinuous, and continuous crutching units. A continuous slurry preparation system is shown in Fig. 26.20. Solids levels normally are kept as high as possible (64–72%) to reduce the heat requirements and maximize the tower throughput. With the introduction of "soft" (biodegradable) alkylbenzene, the sodium tripolyphosphate used in processing slowly changed from low temperature rise (TR) grades to high TR grades. Low TR sodium tripolyphosphate tends to form more "lumps" and "grit" during hydration from the anhydrous to the hexahydrate phase in the mixing stage than the high TR grade. It is advantageous to maximize sodium tripolyphosphate hexahydrate ($Na_5P_3O_{10} \cdot 6H_2O$) formation, as this water is "bound" and will be retained in the final product. The bound water acts as an inert filler and contributes to the free-flowing characteristics during conveying, storage, and packaging.

TABLE 26.3 Spray-Dried Detergent Formulations

Ingredients (100% Active Basis)	*Phosphate-Built Detergent, % in Formula*	*Non-Phosphate-Built Detergent, % in Formula*
Sodium tripolyphosphate	20–35	0
Sodium carbonate	5–10	10–40
Zeolite A	0	15–25
Sodium alkylbenzene sulfonate	0–15	0–20
Sodium alcohol ethoxylate	0–15	0–20
Sodium alcohol ether sulfate	0–5	0–5
Sodium silicate	3–6	3–6
Sodium sulfate	10–30	10–30
Sodium CMC, optical brightener	1–2	1–2
Enzymes	0–2	0–2
Bleaches	0–8	0–8
Water	3–8	1–8

Fig. 26.20. Continuous slurry preparation system.

After the detergent raw materials are mixed, the slurry is charged to a homogenizer, and any large lumps or gritty particles are desized and screened to prevent plugging of the spray nozzle. The slurry then is charged to a booster pump, deaerated, and sent to a high pressure pump (300–1200 psig) for atomization. Most detergent slurries are atomized into spherical droplets via high pressure nozzles. The nozzles produce a hollow conical pattern with a spray angle of 45 to 90°. The slurry emerges from the nozzle as a film, which is sheared into droplets by the centrifugal force developed by specially grooved cores inserted into the nozzle before the discharge point. A series of spray nozzles (2–25) is placed near the top of the spray tower and angled so that one edge of the desired spray pattern is vertical to the spray tower side walls. Pump pressures of from 300 to 1200 psig may be used to atomize the slurry. Some spray towers use two or three rings of nozzles, depending upon the formulation and drying capacity of the unit. Production rates of detergent spray towers range from 5000 to 80,000 lb/hr. In most cases, the rates are controlled by the slurry feed facilities, solids levels, and drying capacity. Figure 26.21 shows a typical schematic for a spray-drying facility.

Spray towers vary in size from approximately 10 to 35 feet in diameter, and range from 40 to 200 feet in height. A majority of detergent spray towers use countercurrent air flow patterns although some are cocurrent in design. In general, a tower of countercurrent design will produce a higher-density particle than the cocurrent type. In countercurrent systems, the heated gas stream (250–270°C) is uniformly dispersed into the lower section of the tower through a series of plenum chambers or baffles equally spaced around the unit's periphery. The hot gases are cooled as they exhaust from the top of the tower by the drying of atomized detergent droplets upon contact with the hot air stream. Because of the relatively low temperatures (110–140°C) at this point in the tower, the particles do not expand to their full capacity as water evaporates. This lower expansion rate results in a bead with increased bulk density and a thicker shell. If a similar formulation were dried in a cocurrent spray tower, the higher initial air contact temperature (250–370°C) and increased drying rate would increase the evaporative rate and result in a thin-walled bead of lighter density.

In most spray towers, the air flow rate and the temperature are adjusted at the inlet ports to maintain balanced air/temperature patterns. When the atomized slurry is sprayed into the drying chamber, the air flow pattern is disturbed and becomes very turbulent. This

Fig. 26.21. Spray-drying facility; the detergent slurry feed can come from a plant like the one in Fig. 26.20.

effect can cause localized temperature gradient zones ("hot or cold spots") and impact the particle drying characteristics. For example, it is possible to have both over and underdried particles discharged at the same time from the tower because of maldistribution of the air flow in the tower. Adjustment of the air flow patterns, as well as sending a balanced spray into the tower, can help counteract short-circuiting. Increased impingement of the product on tower walls is another negative effect of excessive turbulence, and of excessive spray nozzle velocity. A "ring" of product may form approximately 10 to 20 feet below the spray nozzles and build in size until it cracks or scales off the walls. Ring-removal methods such as periodic cleaning using air or water lances, vibration or sonic shocks, or mechanical cleaning rings sometimes are used to prevent a buildup that could result in blockage of the tower discharge.

After drying in the spray tower, the detergent granules are conveyed to storage silos prior to carton packaging. In some processes, product storage is preceded by fluid-bed conditioning.

The product from the tower is very warm (75–120°C) and typically contains free surface moisture levels of 1 to 3 percent. Conditioning is employed to prevent caking or product flow problems due to either the high temperature or excessive free moisture content. Following conditioning, perfume and other heat-sensitive ingredients (enzymes, surfactants, bleaches) may be added to the product. Post-addition is accomplished by metering the additives into the spray-dried base beads in a rotary drum or other mixing device. The post-additions normally comprise 2 to 15 percent of the total detergent product. An excellent review of major spray-drying patents was provided by Sittig.[43]

Agglomeration Processing

Agglomeration is a technique of binding a mixture of granular and/or dry powdered nonuniform raw materials into an attractive granular form with enhanced aesthetic and flow characteristics. The majority of granular automatic-dishwashing compounds (ADWCs) and some laundry products manufactured today use agglomeration processes. The advantages of agglomeration are centered around low capital costs relative to spray towers and reduced energy consumption. Also, agglomeration generally yields a higher-density product, which can result in reduced packaging costs. The primary disadvantage of agglomeration processing is a limitation in surfactant loading due to raw material absorption characteristics, particularly in laundry applications. As energy costs assume increasing importance and higher-density products gain wider acceptance, agglomeration or combinations of agglomeration and spray-drying technologies will find wider use.

The typical process steps for formulating a laundry detergent or ADWC product usually are defined in the laboratory prior to continuous plant processing, but they generally follow the scheme shown in Fig. 26.22. Frequently the dry raw materials are premixed and then metered into the agglomeration equipment. In the agglomerator, the dry raw material surface is exposed to atomized sprays of surfactant or silicates to begin building the agglomerated particle. Following addition of liquids to the dry raw materials, the product has a wet or pasty texture and requires conditioning to remove excess water prior to storage or packaging. Conditioned agglomerates are screened to remove oversize particles, and then are blended with additives such as chlorinated isocyanurates, bleaches, enzymes, dyes, or perfumes that would not be stable under agglomeration conditions.

Numerous equipment types can be used for agglomeration, including horizontal ribbon and vertical mixers, rotary drums, zig-zag mixers, and pan agglomerators. Horizontal ribbon mixers (Fig. 26.23) are used primarily for "dry mix" formulating, where only limited agglomeration occurs. They are comprised of a U-shaped trough and a ribbon or paddle-type agitator. Liquids are sprayed onto the surface of the dry material bed and slowly dispersed into the mix. Mixer operation can be either batch or continuous. Vertical mixers have become more commonly used due to

Fig. 26.22. Unit processes in agglomeration.

Fig. 26.23. Batch or continuous ribbon mixer.

low space requirements and high efficency of operations. The vertical mixer (Fig. 26.24) utilizes high speed agitation to intimately mix and uniformly blend liquids and dry raws prior to discharge. The contact time is reduced to less than 5 seconds. The rotary drum units generally contain baffles that "roll" or "lift" the product for uniform dispersion of liquid feeds. Pan and zig-zag agglomerators generally are used for limited, specialty product manufacture.

The vertical and the rotary drum agglomeration equipment finds most widespread use in the detergent industry becuase of improved agglomeration, increased liquid feed dispersion levels, and optimized product uniformity. In addition to agglomeration of detergents and automatic dishwashing compounds, these units can be used for hydration of sodium tripolyphosphates, dry neutralization of sulfonic acids, post-addition of raw materials and fines to spray-dried products, and particle encapsulation. An excellent review of detergent agglomeration has been provided by Dolan.[44]

Liquid Detergent Processing

Liquid heavy duty laundry products accounted for approximately 40 percent of the 1990 U.S. heavy duty laundry detergent market. The

Fig. 26.24. Continuous in-line vertical agglomerator.

growth in heavy duty liquids began in 1974 as the result of phosphate-banning legislation. Unbuilt liquids based on nonionic and anionic/nonionic active combinations gradually increased in volume for various reasons including convenience, solubility, and improved performance of these types versus nonphosphated powdered detergents. In the mid-1980s the introduction of higher-performance liquids accelerated growth in this segment, with the market share essentially doubling from 1984 to 1990.[45]

Liquid products are approximately 50 to 60 percent water, with the remainder being a combination of surfactants, builder, foam regulators, hydrotropes, anti-redeposition agents, whitening agents, corrosion inhibitors, colorants, and perfume. Several formulations also contain antistatic and fabric-softening ingredients. Liquid systems require careful selection and blending of raw materials to achieve a stable product. Special attention is necessary for the following items:

- *Viscosity.* The product must be pourable and must retain the same viscosity characteristics from batch to batch. This variable can be controlled by proper solvent or hydrotrope selection.
- *Clear-cloud point.* The composition should have sufficient solubility to prevent "hazing" or separation of product when subjected to storage in cool temperatures.
- *Freeze-thaw stability.* The formulation must be compounded to prevent phase separation or solidification at freezing temperatures.

Many light duty liquid detergents are formulated with linear alkylbenzene sulfonate. Heavy duty liquid detergents often use a combination of actives such as LAS, sodium alcohol ether sulfate, alcohol ethoxylates, and olefin sulfonates. The manufacture of light and heavy duty formulations requires relatively basic equipment for blending of the various ingredients. The raw materials are added batchwise to a large mixing vessel or proportionated to in-line mixers before being directly fed from the mixer to the filling lines.

In 1986, liquid automatic-dishwashing detergents were introduced. These products used the principle of thixotropic slurries to address product form issues such as machine cup dispensing rates and product physical separation. The slurries were generally homogenous in nature and sufficiently viscous to prevent premature dispensing into the dishwasher prior to the wash cycle. Although the market share had grown to approximately one-third of the automatic-dishwashing market by 1988, product issues of cleaning performance, unused product remaining in the package, and unwanted bleaching of kitchen fabrics now have decreased product market share to less than 20 percent. Unique ingredients in this product are bentonite clays, which are used to thicken formulations and provide their characteristic viscosity. High-shear mixing equipment such as homogenizers is used for this application in batch or continuous stirred tank systems. Product types and processing techniques are covered in patents.[46–49]

Detergent Trends

To date, no new detergent builders have been developed that have the functional

characteristics and cost advantages of phosphates. A variety of builders (previously discussed) presently are being formulated in nonphosphate-legislated detergent products, but these salts are not as effective as phosphates. This performance differential is being compensated by increased use levels of synthetic surfactant and other detergent additives.

The major worldwide detergent product trend is introduction and sale of powdered detergent concentrates. These laundry products, which achieved widespread market dominance in Japan in 1988, have captured sizable market acceptance in Europe, were test-marketed in the United States, and since then have been claiming an increasing share of the market in the United States. These dense products permit reduced use levels of one-third to one-half cup per load and provide a significant reduction in product packaging, warehousing, and shipping costs. The product type requires process changes that utilize combinations of spray drying, agglomeration, and sometimes granulation technologies.[50–52]

Although growth in liquids appears to have plateaued, liquid detergent formulations are expected to maintain their position in the detergent market. Current and future energy considerations as well as water conservation suggest that a movement toward reduced water temperatures and water consumption will occur. These factors, coupled with high consumer performance expectations, will continue to challenge the detergent industry during the coming decade.

REFERENCES

1. *Soaps and Detergents*, The Soap and Detergent Association, New York, 1981.
2. Gupta, S., "Chemistry, Chemical and Physical Properties and Raw Materials," in *Soap Technology for the 1990's*, L. Spitz (Ed.), pp. 48–93, American Oil Chemists Society, 1990.
3. By courtesy of G. Mazzoni S.p.A., Busto Arsizio, Italy. See also Spitz, L., *J. Am. Oil Chem. Soc.*, **45**, 423 (1967).
4. For general references, see Dieckelmann, G., and Heinz, H. J., *The Basics of Industrial Oleochemistry*, Peter Pomp GmbH, Essen, W. Germany, 1988; Woollatt, E., *The Manufacture of Soaps, Other Detergents and Glycerin*, Ellis Horwood, Sussex, England, 1985.
5. By courtesy of Alfa-Laval, Tumba, Sweden.
6. By courtesy of Lurgi GmbH, Frankfurt am Main, Germany.
7. Sonntag, N. O. V., *J. Am. Oil Chem. Soc.*, **56**, 729A, 861A (1979); Markley, K. S., *Fatty Acids*, 2nd ed., Interscience, New York, 1967.
8. Gosewinkel, L., and Knuth, M., *Fett Wiss. Technol.*, **90**, 155 (1988); Stoiculescu, P., and Svet, V., *Ind. Aliment. (Bucharest)*, **22**, 320 (1971).
9. Leshchenko, P. S., *Maslo-Zhir. Prom.*, **11**, 21 (1973).
10. Linfield, W. M., Barauskas, R. A., Silvieri, L., Serota, S., and Stevenson, R. W., Sr., *J. Am. Oil Chem. Soc.*, **61**, 191 (1984); Linfield, W. M., O'Brien, D. J., Serota, S., and Barauskas, R. A., *J. Am. Oil Chem. Soc.*, **61**, 1067 (1984).
11. Brady, C. L., Metcalfe, L., Slaboszewski, S., and Frank, D. *J. Am. Oil Chem. Soc.*, **65**, 917 (1988); Kwon, D. Y., and Rhee, J. S., *Korean J. Chem. Eng.*, **1**, 153 (1984).
12. Park, Y. K., Pastore, G. M., and de Almeida, M. M., *J. Am. Oil Chem. Soc.*, **65**, 252 (1988); Holmberg, K., and Osterb, E., *J. Am. Oil Chem. Soc.*, **65**, 1544 (1988).
13. Mange, H. K., and Vaidya, S. D., *J. Oil Technol. Assoc. India*, **11**, 73 (1979). Vaidya, S. D., Subrahmanyan, V. V. R., and Kane, J. G., *Indian J. Technol.*, **13**, 528 (1975).
14. Berger, R., and McPherson, W., *J. Am. Oil Chem. Soc.*, **56**, 743A (1979); Jach, K. W, and Stage, H., *Fett Wiss. Technol.*, **90**, 501 (1988).
15. Stage, H. *J. Am. Oil Chem. Soc.*, **61**, 204 (1984); Jach, K. W., and Stage, H., *Fett Wiss Technol.*, **90**, 501 (1988).
16. Haradsson, G., *J. Am. Oil Chem. Soc.*, **61**, 219 (1984); Zilch, K. T., *J. Am. Oil Chem. Soc.*, **56**, 739A (1979).
17. Luddy, F. E., *J. Am. Oil Chem. Soc.*, **56**, 759A (1979).
18. Haupt, D. E. *Soap/Cosmetics/Chem. Spec.*, **60**, 42 (1984).
19. Leikham, J. W., *J. Am. Oil Chem. Soc.*, 89 (1988).
20. Swisher, R. D., *Surfactant Biodegradation*, Surfactant Science Series, Vol. 3, Marcel Dekker, New York, 1970.
21. Huddleston, R. L., and Nielsen, A. M., *Household & Personal Products Industry*, p. 72 (1979).
22. Gledhill, W. E., Huddleston, R. L., Kravetz, L., Nielsen, A. M., Sedlak, R. I., Vashon, R. D., *Tenside*, **26**, 276 (989).
23. Schoberl, P., *Tenside*, **26**, 86 (1989).

24. Brooks, R. J., et al., U.S. Patent 3,257,715, June 21, 1966.
25. Brooks, R. J., et al., U.S. Patent 3,259,645, July 5, 1966.
26. Ballestra, M., U.S. Patent 3,180,699, Apr. 27, 1965.
27. Lantini, A., U.S. Patent 3,931,273, Jan. 6, 1976.
28. Hennig, H., et al., U.S. Patent 2,802,026, Aug. 6, 1957.
29. Logan, R. S., U.S. Patent 3,188,183, June 8, 1965.
30. Brooks, R. J., et al., U.S. Patent 4,113,428, Sept. 12, 1978.
31. Brooks, B., U.S. Patent, 4,185,030, Jan. 22, 1980.
32. Zallestra, M., U.S. Patent 3,180,699, Apr. 27, 1965.
33. Sumner, C. A., U.S. Patent 3,597,361, Aug. 3, 1971.
34. Frank, R. A., et al., U.S. Patent 3,867,316, Feb. 18, 1975.
35. Schick, M. J., *Nonionic Surfactants*, Vol. I, Marcel Dekker, Inc., New York, 1987.
36. Jungerman, E., *Cationic Surfactants*, Marcel Dekker, Inc., New York, 1970.
37. *Mc'Cutcheon's, Detergent and Emulsifiers*, North American ed., MC Publishing Company, Glen Rock, NJ, 1988.
38. Swartz, A. M., and Perry, J. W., *Surface Active Agents*, Vol. 1, John Wiley & Sons, New York, 1949.
39. Jungermann, E., *Cationic Surfactants*, Marcel Dekker, Inc., New York, 1979.
40. VanWazer, *Phosphorus And Its Compounds*, Vol. II, John Wiley & Sons, New York, 1961.
41. Sorgenfrei, M. E., "Increased Use of STPP," *Soap/Cosmetics/Chem. Spec.*, **3** (2), 60 (Feb. 1980).
42. Findley, W. R., *JAOCS*, **65** (4) (1988).
43. Sittig, M., "Detergent Mfg. Including Zeolite Builders," *Chem. Tech. Review*, No. 128, Noyes Data Corporation, NJ (1979).
44. Dolan, M. J., "Detergent Agglomeration Technology," *Soap and Chemical Specialties* (Mar./Apr. 1988).
45. Cannon, D. R., *Chemical Week*, p. 22 (Jan. 21, 1987).
46. Lancz, A. J., GB 2,163,448, Aug. 13, 1985.
47. Julemont, M., DE 3,325,503, Nov. 29, 1984.
48. Lai, K., GB 2,164,350, Aug. 13, 1985.
49. Colarusso, R. M., GB 2,163,447, Aug. 13, 1985.
50. Gangwisch, W., et al., U.S. Patent 4,264,464, Nov. 11, 1989.
51. Nakamura, R., and Sai, F., SHO 64[1989]-20298, Nov. 24, 1989.
52. Donaldson, R., et al., European Patent 340,013, Nov. 2, 1989.

27

Pigments, Paints, Polymer Coatings, Lacquers, and Printing Inks

R. A. Ryntz* and V. E. Gunn**

INTRODUCTION

Change is constant in the coatings market. As mergers, acquisitions, and partnerships take shape, consolidation remains prominent. The 80/20 rule (20% of the firms accounting for 80% of the business) takes effect as the need for regulatory and environmental compliance continues to plague the market.

The paint and coatings market garnered $35 billion in sales (Fig. 27.1) in 1990. The major participants include ICI (London), BASF (Ludwigshafen), Hoechst (Frankfurt), Du Pont (Wilmington, DE), PPG Industries (Pittsburgh, PA), Akzo (Arnhem, the Netherlands), and Nippon Paints (Tokyo). Smaller companies, being forced out of the market because they lack the capital required to meet compliance issues, are forced into consolidations. Some key acquisitions made in 1989 in the coatings industry are listed in Table 27.1. Other consolidation actions are coming from unexpected sources. For example, UB Group (Bangalore, India), through its subsidiary Jenson and Nicholson (Dagenham, U.K.), bought two British paint companies, Cementone-Beaver and Windeck Paints, to take a 50 percent interest in a third, Buckingham Coatings—the first time an Asia-based paint company has made an acquisition in Europe.[1]

Fig. 27.1. Key worldwide coatings markets (*Source: Kline and Co., Fairfield, NJ. Chemical Week estimates.*)

*Akzo Coatings, Inc., Troy, MI.
**P.O. Box 46315, Mt. Clemens, MI.

TABLE 27.1 Key 1989 Acquisitions in the Coatings Industry[2]

Company Location	*Acquisition*	*Price (in million $)*
• Courtaulds (London)	Products Research Chemicals	282
• Akzo (Arnhem, The Netherlands)	Reliance Universal	275
• Sherwin–Williams (Cleveland)	Whittaker's coatings and chemicals division	147
• Dainippon (Tokyo)	NL Industries/Spencer Kellog	86
• Nippon Oil and Fats (Tokyo)	U.S. Paint (Grow Group Subsidiary)	25
• Lilly Industrial Coatings (Indianapolis)	Jamestown Coatings (Essex Spec. Prods. subsid.)	NA
• Valspar (Minneapolis)	McCloskey	NA
• PPG (Pittsburgh)	Olympic; Lucite (Clorox)	130
• RPM (Medina, OH)	Kop-Coat	48

Source: Kline and Co. (Fairfield, NJ).

TABLE 27.2 U.S. Sales of Paints and Coatings (millions of dollars)[1]

	1987	*1988*	*1989*	*1990**	*1992**	*1995**
Architectural	4,245	4,372	4,705	4,990	5,820	7,420
Product finishes	3,782	4,063	4,371	4,666	5,485	7,090
Special-purpose coatings	3,140	2,281	2,484	2,680	3,160	4,070
Miscellaneous allied paint products	995	1,036	1,070	1,100	1,240	1,500
Total	11,161	11,750	12,639	13,436	15,705	20,080

*Projected
Source: Leading Edge Reports (Cleveland).

TABLE 27.3 Western European Paints and Coatings Production (millions of dollars)[1]

Country	*1981*	*1985*	*1987*	*1990*
West Germany	3,039	3,583	3,833	4,312
France	2,052	2,111	2,299	2,585
United Kingdom	2,113	2,028	1,875	2,111
Italy	2,217	1,585	1,771	1,993
Netherlands	681	735	798	898
Spain	642	640	668	751
Belgium	392	439	484	544
Sweden	434	465	448	504
Switzerland	282	345	390	439
Denmark	295	379	371	417
Austria	311	325	311	350
Finland	291	271	262	295
Norway	230	278	249	280
Portugal	194	173	191	214

Source: Information Research Ltd. (London) and *CW* estimates.

In the United States the major players of over 1500 producers include PPG, Du Pont, Sherwin-Williams, BASF, Akzo, and Glidden/ICI. The market in 1989 consisted of $12.6 billion in sales (see Table 27.2), and was expected to grow to over $20 billion by 1995.

Restructuring also continues in Western Europe's $16 billion/year paint and coatings market, which decreased from some 1800 players in 1980 to approximately 1500 in 1990 (see Table 27.3).

The major coatings markets in the United States consist of architectural coatings (consumer and decorative), product finishes (maintenance/protective coatings for bridges, industrial machinery, etc.), special purpose coatings such as automotive finishes (DEM and refinish), and miscellaneous allied paint products (e.g., for appliances, containers, and furniture and related materials) (Table 27.2).

This market is being plagued by environmental constraints imposed by local, state, and federal governments. The key concerns of government regulators and consumers include:[3]

1. Treating the earth as one biosphere, having a finite capacity for cleansing itself of toxins.

2. Cleaning air from pollutants that are killing lakes and causing a dramatic increase in respiratory illnesses.
3. Ensuring that smog-enshrouded cities do not become industrial nightmares.
4. Making the agricultural and livestock products that people eat safe for consumption over the long term.

These concerns have led to legislation over the last several years that causes manufacturers to place a greater emphasis on research to meet and exceed government and consumer demands for clean air.

VOC REGULATIONS

The environmental demands have led to a nationwide program to limit volatile organic compounds (VOCs). The volatile organic compounds of a paint are calculated as pounds of solvents per gallon of paint. Prior to 1970, the VOC content of most paints was well above 5 lb/gal. Current major industrial paints now are limited to approximately 3.5 lb/gal of VOC because of imposed environmental pressures. Stricter legislation in the future, however, will place requirements of less than 3.0 lb/gal of VOC on most markets.

The issue of air quality compliance is complicated by different regulations in various states and regional areas. Although the federal government (through the U.S. EPA) has set minimum national standards, states and designated regional areas can adopt their own compliance rules.

All states and regions tend to restrict the solvent content in coatings. The following summaries indicate the wide variation in compliance requirements throughout the country.[4]

West

California. This is where the regulations began. The regulatory activities of the South Coast Air Quality Management District (SCAQMD) have set a pattern that many other regulatory boards have chosen to follow. The district restricts application equipment and VOC content. The VOC requirements are continuously changing; so no attempt was made to list them here.

The SCAQMD in Southern California has mapped out a three-tier plan to bring the district into clean-air compliance. The plan will be revised periodically based on reports from over 40 monitoring stations throughout 13,350 square miles including Los Angeles, Orange, and Riverside counties and the urban portion of San Bernardino County. VOCs will not be monitored directly. Rather, ozone, which is formed when hydrocarbons (VOCs) react with nitrogen oxides in sunlight, will be measured for the presence of VOCs. The federal clean-air standard for ozone is not to exceed 0.12 ppm more than one day a year.

Tier I of the three-tier plan involves 22 rules to be adopted by December 31, 1993, to restrict the emission of VOCs from coatings. Control methods include the application of low-solvent coatings, use of high transfer efficiency techniques to apply coatings, and "to a certain extent" the use of add-on control devices (see Table 27.4).

Tier II (to be implemented by 2000) calls for reducing the remaining VOC emissions from coatings by 50 percent. The reductions will be achieved by further tightening of coatings VOC content and application transfer efficiency requirements.

Tier III (to be implemented by Dec 31, 2007) calls for applying new technology to improve waterborne, UV-cure, and two-component coatings and prohibiting "certain high-emitting uncontrollable coating processes."

Washington. No rule prohibits any method of spray finishing. VOC guidelines are determined by subtracting the amount of solids from the gross weight of the material used and determining whether the amount of solvent used exceeds the acceptable limits.

The maximum VOC limits (lb/gal) are: clearcoat, 4.3; color, 3.5; performance coatings, 3.5; all other coatings, 3.0. Metric limits (g/L) are defined for aerospace: primers, 650; topcoats, 600 (equivalent to 5.4 and 5.0 lb/gal, respectively).

TABLE 27.4 Application Equipment Requirements of SCAQMD Clean Air Act[4]

Application			*Rule*			**	***			
Equipment	1104	1106	1107	1113	1115	1124	1125	1136	1145	1151
• Dip			×			×	×	×		
• Electrostatic application			×			×	×	×		×
• Flow coat			×			×	×	×		
• Hand-application methods						×	×	×		
• HVLP* spray			×			×	×	×		×
• Roll coater			×			×	×	×		
• Other methods that are proven			×			×	×	×		×
• 65% transfer efficient										
• No specified application method	×	×		×	×				×	

*High volume low pressure application equipment.
**After 1/1/92.
***Or until 7/1/94, use air-assisted airless spray and coatings that meet the 7/1/96 limits.

Southwest

Texas. Here, regulators restrict emissions rather than equipment. Emissions are limited across the board for certain industries. Limits (lb/gal) are specified for particular coatings in other industries; for example: large appliances, 2.8; furniture, 3.0; coil coating, 2.6; automotive original equipment manufacturer (OEM)—primers, 1.2 if submersed and 2.8 if sprayed, topcoats, 2.8, and repair, 4.8; miscellaneous metal—color, 3.5, top clearcoat, 4.3, and all others, 3.0.

South Carolina. VOC limits (lb/gal) include: metal furniture, 3.0; appliances, 2.8; miscellaneous metal—color, 3.5, clearcoat, 4.3, and all others, 3.0; coil coating, 2.6.

North Carolina. The rules are applicable only in areas designated as nonattainment zones. If air pollution control equipment is not used, emission standards are based on the weight of solids used. High-solids coatings are automatically compliant regardless of spray equipment used. The limits (lb/gal) for automotive shops with air pollution control equipment include: primers, 1.4 if high-solids, 1.2 if low-solids; topcoats, 4.5 if high-solids, 2.8 if low-solids; repair, 13.8 if high-solids, 4.8 if low solids. For automotive shops without air pollution control equipment limits are: primers, 1.2; topcoats, 2.8; repair, 4.8. Air pollution control equipment is not concisely defined. Other limits include: metal furniture, 5.1; appliances, 2.8.

Georgia. VOC emissions are limited (lb/gal) by industry and coatings used: automotive—primers, 1.2 if dipped and 2.8 if sprayed, topcoat, 2.8, and repair, 4.8; metal furniture, 3.0; appliance, 2.8; and miscellaneous metal—performance coatings, 3.5, high-performance coatings, 6.2, clearcoat, 4.3, and all others, 3.0; coil coating, 2.6.

Midwest

Indiana. Rules in Indiana are among the most complex in the nation. They require a baseline TE (transfer efficiency), specify coatings formulations, and limit VOC emissions (lb/gal) for automotive, metal furniture, and large appliances, but then only specify acceptable equipment choices for wood. Limits are: automotive—30 percent TE at 62 percent solids; metal furniture—60 percent TE at 59.2 percent solids; large appliances—60 percent TE at 62 percent solids. No TE test procedure is defined for the following VOC limits: automotive—primer, 1.9, topcoat, 2.8, and

repair, 4.8; metal furniture, 3.0; large appliances, 2.8; miscellaneous metal—color, 3.5, specialty coats, 3.5, clear-coats, 4.3, coil coating, 2.6, and all others, 3.0. In the wood industry the state departs from these guidelines and identifies acceptable equipment choices: airless, heated airless, and air-assisted airless spray; manual and automatic electrostatic spray; and rollers, brushes, or wipes. Significantly, no conventional air spray method, including HVLP, is recognized as compliant.

Illinois. A 30 percent minimum TE is required for automotive. Automotive VOC limits (lb/gal) are: primer/surface, 2.8; topcoat, 2.8; repair, 4.8. The rules allow higher VOC limits for automotive topcoats if the finishing method can be demonstrated to achieve a TE of over 65 percent. In such cases topcoats may be used with a VOC of up to 3.6. For the wood industry a minimum TE of 65 percent is required with the following VOC limits: washes, 6.5; stains, 4.7; topstains, 6.6; sealers, 5.6; clearcoats, 5.6. No test procedure for TE has been defined.

Michigan. Althhough VOC guidelines are suggested, they are not stated clearly. In general the finisher is expected to achieve minimal VOC levels when using metallic topcoats of waterborne coatings and performing touchup or repair work. Although conventional air spray is generally acceptable, some areas of Michigan require electrostatic systems with a TE of at least 40 percent and define VOC limitations.

East

New Jersey. A minimum of 65 percent TE is required for all industries except wood. Shops using waterborne coatings are automatically compliant regardless of TE. VOC limits (lb/gal) for all others are: automotive—primers, 2.8, topcoats, 2.8, specialty coats, 5.0, and repair, 4.8; automotive refinishing—basecoats, 6.0, clearcoats, 4.4, and all others, 5.0; coil coating, 2.6; miscellaneous metals—color, 4.3, performance coatings, 3.5, and all others, 3.0; leather, 5.8; urethane users, 3.8. Wood industry rules specify equipment by the shop size. Wood shops using less than 50 tons of coatings a year must achieve a minimum TE of 40 percent. Approved spray-finishing methods include airless, heated airless, and air-assisted airless. Wood shops that use 50 or more tons of coatings a year are required to achieve a TE of 65 percent or higher. Approved spray-finishing methods for these shops include airless, heated airless, air-assisted airless, electrostatic spray, or flatline finishing. VOC limits (lb/gal) are: washcoat, 6.1; stain, 4.7; pigment, 5.0; sealer, 5.6; clearcoat, 5.6.

Pennsylvania. The rules specify only that wood cabinet and furniture finishers must use air-assisted airless or electrostatic spray.

New York. New York follows the same rules as Pennsylvania.

Ohio. Ohio is still adopting VOC limitations and will require a TE of at least 30 percent.

HAZARDOUS WASTE REGULATIONS

Included in regulatory compliance is the issue of hazardous wastes. On March 29, 1990, the U.S. EPA promulgated a final rule revising the procedure used to determine whether wastes are hazardous because of their potential to leach significant concentrations of toxic constituents to landfills. The extraction procedure (EP) toxicity characteristic procedure was replaced with the Toxicity Characteristic (TC), utilizing the Toxicity Characteristic Leaching Procedure (TCLP).[5] The U.S. EPA also added to the list 25 organic chemicals commonly used in surface-finishing operations (i.e., methyl ethyl ketone, tetrachloroethylene, etc.)

Although the maximum allowable concentrations of the regulated eight metals (and pesticides, herbicides, and insecticides) do not change from those listed under the EP toxicity, the new TCLP analytical procedure may make it more difficult for listed wastes to stay under the allowable concentrations (Table 27.5).

TABLE 27.5 Maximum Concentration of Contaminants Allowable from TCLP*,[5]

EPA Haz Waste No.	*Contaminent*	*Regulatory Level (mg/liter)*
D004	Arsenic	5.0
D005	Barium	100.0
D018	Benzene	0.5
D006	Cadmium	1.0
D019	Carbon tetrachloride	0.5
D021	Chlorobenzene	100.0
D022	Chloroform	6.0
D007	Chromium	5.0
D023	*o*-Cresol	200.0
D024	*m*-Cresol	200.0
D025	*p*-Cresol	200.0
D026	Cresol	200.0
D027	1,4-Dichlorobenzene	7.5
D028	1,2-Dichloroethane	0.5
D029	1,1-Dichloroethylene	0.7
D030	2,4 Dinitrotoluene	0.13
D032	Hexachlorobenzene	0.13
D033	Hexachlorobutadiene	0.5
D034	Hexachloroethane	3.0
D008	Lead	5.0
D009	Mercury	0.2
D014	Methoxychlor	10.0
D035	Methyl ethyl ketone	200.0
D036	Nitrobenzene	2.0
D037	Pentachlorophenol	100.0
D010	Selenium	1.0
D011	Silver	5.0
D039	Tetrachloroethylene	0.7
D040	Trichloroethylene	0.5
D041	2,4,5-Trichlorophenol	400.0
D042	2,4,6-Trichlorophenol	2.0
D043	Vinyl chloride	0.2

*Excluding D012–D017 and certain pesticides.

TECHNICAL TRENDS IN COATINGS

Through the years, improved paints have been developed because of competition from environmental and consumer activists. The objective of improving overall quality and reducing emissions of VOCs has led to a wider selection of paint types. The user, therefore, must select a coating type that meets the cost of required application, equipment, energy costs for using, pollution and waste collection expenses, and application and cured performance requirements of the coating.

Table 27.6 lists types of coatings used in industrial and special purpose coatings by technology. Chemark-based forecasts show increased usage of all types of materials in 1996 except for conventional noncompliant coatings, lacquers, and nonaqueous dispersions. Overall, consumption of solvent-based materials is expected to decline from a 74 percent level in 1986 to 63 percent in 1996. This compares to an 85 percent level in 1975.

A solvent-borne coating generally consists of seven major constituents.

- Resin (binder)
- Cross-linker
- Rheology control agent
- Catalyst
- Pigment
- Flow additive
- Solvent

The resin (often referred to as binder) is responsible for the main attributes of the coating when cured with the cross-linker (or

TABLE 27.6 Technology Change in Coatings Types Expected from 1986 to 1996 (Percent of Total Dry Gallons)[6]

Type of Coating	*1986*	*1996*	*Rate of Change*
Solvent-based			
Conventional low solids[a]	42	30	(1.8)
Lacquers	9	5	(3.2)
Two-part catalyzed	16	19	3.0
High solids	6	9	3.5
Nonaqueous dispersion	1	0	—
	74	63	
Water-based			
Emulsion/latexes	8	10	1.7
Water-solubles and dispersions[b]	13	18	4.6
	21	28	
Other			
Powder	4	6	7.2
Radiation curable	1	3	6.6
	5	9	
Total	100	100	

[a]Except lacquers.
[b]Includes electrocoats.
Source: Based on Chemark estimates.

cross-linking agent). The curing, or network formation, often is accelerated through the use of a catalyst. Rheology control agents, often referred to as sag control agents, are added to impart shear thinning behavior when under high shear (i.e., spray application) and rheology or antisagging behavior when under low shear (applied on substrate). Pigments, discussed later in this chapter, are introduced to impart color.

The amount of pigment utilized depends on the color and the hiding power required of the coating. The flow additive often is introduced to relieve surface tension variations between the coating and substrate, to eliminate pinholes or crater formation. Solvents are added as necessary to achieve flow under application conditions.

In order to achieve a VOC-compliant coating, one must consider the molecular weight of the main binder utilized in the coating formulation, as well as the effect of additives on the final viscosity. After all, application viscosity, which in turn is affected by the rheology of the coating constituents, is most often controlled by freely adding additional solvent to achieve necessary flow and leveling upon application. This in turn affects the VOC emission of the coating.

If we look at the molecular weight dependency of resins utilized in the various coating types, it is readily evident why VOC compliance is more easily achieved in radiation-curable and powder coatings. The molecular weight of nonaqueous dispersion resins, for example, exceeding 10^5 g/mole, results in very viscous materials, whereas the typical molecular weight of a radiation cure oligomer ranges from 10^2 to 10^3 g/mole, for a very low viscosity material, which requires little if any additional solvent to attain application specifications.

If we consider the evolution of coating technology over the last 70 years (see Table 27.7) we can see that by the mid-1950s waterborne and higher solids coatings already were under development. Powder coatings, 100 percent reactive liquid coatings (electron/beam and UV cure), electrocoatings, aqueous or waterborne (water-base) coatings, and high solids coatings can all be considered as novel "nonpolluting" VOC-compliant coating systems.

TABLE 27.7 Evolution of Coatings Development*

1923	Low viscosity nitrocellouse
1923	Alkyds
1924	Titanium dioxide
1928	Phenolic resins (oil-soluble)
1930	Urea formaldehyde–alkyd mixtures
1933	Vinyl chloride copolymer
1934	Oil-based emulsion paints
1936	Thermoset acrylics
1937	Polyurethanes
1939	Melamine formaldehyde–alkyl mixtures
1944	Silicon-based paints
1947	Epoxy resins
1950	Polyvinylacetate and acrylic emulsions
1955	Powder coatings
1958	Latex house paint
1960	Water-based coatings
1962	Anodic electrodeposition
1963	UV and electron beam cure
1971	Cathodic electrodeposition
1974	Clear over color topcoats
1989	Supercritical CO_2

*J. Coating Technology, **62** (780), 83 (1990).

Powder Coatings

Powder coatings consumption in 1985 was 70 MM lb valued at $163 million. In 1990, 118 MM lb were consumed at a value of $300 million.[1,7] Today, they are the fastest growing finishing technology in North America with a growth rate (in pounds sold) of 15 to 20 percent per year until at least 1995.

The major impetus for growth can be attributed to the virtual "pollution-free" finishing system as well as the dramatic leap forward in materials and equipment technology.

Powder coating resin systems (see Fig. 27.2) have been developed to meet specific needs of the metal finishing operations. Epoxies were utilized almost exclusively during the early years of thermosetting powder coatings and continue to be utilized in a range of applications. However, because of the poor outdoor

Fig. 27.2. Powder coating production by resin type (*Bocchi, G.,* Modern Paint and Coatings, *p. 36, Nov. 1989.*)[8]

durability of epoxy resins, acrylics and polyester–urethane hybrids are increasing in use (see Table 27.8).

Powder coating manufacturers continue to work on perfecting flow and leveling of the coatings, which often are prone to orange peel. The other major disadvantage of thermosetting powder coatings is the high temperature (300–350°F) required to cure these systems. Research focuses on lowering the temperature of cure.

Electron Beam (EB) and Ultraviolet (UV) Curable Coatings

Radiation-curable coatings were developed by W. J. Burlant of the Ford Motor Company in 1962. Acceptance of finishes designed for radiation curing (high solids materials formulated to be cured by UV and EB systems) has been slow, and current consumption is low (about $250 million).

The major advantages of radiation-curable coatings are speed of cure, reduced emissions, 100 percent solids systems, and elimination of ovens or other sources of thermal energy. However, there also are disadvantages: it is necessary to formulate clear or transparent coatings to allow UV light penetration; the appearance of these coatings is somewhat lower in quality than that of conventional finishes; UV coatings are toxic because of the chemical nature of the monomers that compose the coating; and uniform subtrates are necessary for UV radiation.

The UV coating usually is comprised of unsaturated monomers that polymerize with UV light (in the presence of a photoinitiator, which absorbs radiation and begins polymerization). Unsaturated polyesters copolymerized with vinyl monomers or multifunctional acrylates often are utilized. Most recently, however, vinyl ether monomers for cationic curing with epoxies have shown

TABLE 27.8 Letterpress and Lithographic Inks—Dry Mechanisms[12]

Type Ink or Process	*Principal Ink Vehicle*	*Typical Substrates*	*Principal Ink Drying Mechanism*	*Remarks*
Letterpress W	Mineral oil	Newsprint	A	Nonvolatile mineral oil
Letterpress S	Oleoresinous and solvent	Papers, boards	A, O	High-boiling range solvent
Letterpress S	Resin and glycol	Papers, boards	P	High-boiling glycol
Letterpress S	Polyfunctional (acrylated) compounds	Papers, boards, foils,	R	No solvent Ultraviolet curing
Offset Litho W	Mineral oil and resin and solvent	Newsprint	A	High-boiling solvent
Offset Litho S	Oleoresinous and solvent	Papers, boards	A, O	High-boiling solvent
Offset Litho S	Oleoresinous and solvent	Tinplate	O,E (stoving)	High-boiling solvent
Offset Litho S	Polyfunctional (acrylated) compounds	Papers, boards, foils, tinplate	R	No solvent Ultraviolet curing
Offset Litho W	Resin/solvent	Papers	A, E (Heat-set)	Gas flame hot air High-voiling solvent

A = Absorption O = Autoxidative/polymerization P = Precipitation/gelation R = Photochemical/polymerization, ultraviolet cure S = Sheet-fed W = Web-fed

promise due to reduced shrinkage of cured films and improved adhesion to substrates.

Major uses for UV curing (room temperature) include use on flatstock-fillers in particle boards and hardwood flooring, can protective coatings, high-gloss (no-wax) coatings for floor tiles, wood finishes, and other specialty uses. Future growth of 10 to 15 percent per year is anticipated.

Current Automotive Coating Trends

Until 1965 automotive topcoats consisted of nothing more than acrylic lacquer technology. As technology improved to solvent-borne nonaqueous dispersions in the mid-1960s and medium solids thermoset coatings in the 1970s, two competing factors were becoming more important: control of emissions (VOC) and the quality of the appearance of the topcoat. Today, the quality of the topcoat includes chip resistance, environmental etch resistance, and metallic glamor; so virtually all automotive procedures (except in truck plants) now utilize basecoat/clearcoat technology. The basecoat is the pigmented film that is responsible for providing color. It is covered, in a wet-on-wet spray process, with a clearcoat that provides chemical resistance, UV resistance, chip resistance, and distinctness of image (DOI) (often referred to as the "wet look").

The total automotive coating composite, however, consists of more than just the basecoat/clearcoat. The metal sheet stock that is formed into the automotive car body is treated with a corrosion-resistant primer applied by cathodic electrophoresis. This coating, often referred to as "ECOAT" or "uniprime," is a hydroxy functional epoxy/blocked isocyanate dispersion that deposits onto the cathode of an electrolytic cell at a voltage of between 250 and 425 volts. Film thickness, which is dependent upon the residence time, temperature, and coating voltage of the electrocoat bath (electrolytic cell), varies between 0.8 and 1.2 mils (1 mil = 25 microns). The curing temperature required to deblock the isocyanate and yield a cross-linked network varies from 275°F to 375°F for 20 minutes. The function of this coating is to provide a corrosion-resistant film to the metal sheet stock.

A primer surfacer, composed most often of a polyester to be cross-linked with a blocked isocyanate or melamine cross-linker, is applied next at a dry film thickness of approximately 1.0 mil by conventional spray application. The function of this coating is to provide filling and leveling properties as well as some stone-chip resistance. Cure conditions range from 250°F to 325°F for 20 minutes.

The solvent-borne basecoat, consisting of polyester or acrylic binder to be cross-linked with melamine or isocyanate, is applied at a dry film thickness of between 0.6 and 1.5 mils, depending upon the color. Poorer-hiding colors, such as red and white straight shades, require the higher film builds, whereas metallics and darker colors are applied at the lower film builds. The solvent-borne basecoat is allowed to flash-off solvent for approximately 3 to 5 minutes at ambient conditions prior to application of the solvent-borne clearcoat. This process is referred to as a wet-on-wet application because the basecoat is not in its cured or cross-linked state.

Optionally, a waterborne basecoat can be utilized to achieve better aluminum orientation in metallic films. Several papers[9,10] have described metallic orientation and the advantages of waterborne basecoats; so this concept will not be thoroughly discussed here.

The waterborne basecoat consists of a urethane or an acrylic dispersion with passivated aluminum for metallic glamor. The aluminum must be passivated to prevent reaction with water and the evolution of hydrogen gas. Passivating agents can include polymers, hexavalent chromium, or phosphate esters. The basecoat is prebaked (with infrared or thermal ovens) to release the majority of the water prior to clearcoat application.

The solvent-borne clearcoat is spray-applied over the basecoat at a dry film thickness of 1.8 to 2.0 mils. The clearcoat, responsible for UV protection, chemical resistance, chip resistance, and so on, is most often composed of a hydroxy functional acrylic binder together with a melamine or an isocyanate cross-linker. Ultraviolet light absorbers and hindered amine light stabilizers also are added to achieve better free-radical oxidative degradation resistance when the coating is exposed to sunlight.

The clearcoat is referred to as a two-component (2K) clearcoat when isocyanate is used as the cross-linker. The isocyanate must be added to the hydroxy functional acrylic just prior to spray application to prevent premature cross-linking (or gellation). For this reason, the two-part mixing is referred to as 2K. Cure conditions for isocyanate-cured clearcoat range from room temperature (car refinish) to 250°F for 30 minutes.

One component, or 1K, clearcoats are equivalent to 2K clearcoats with the exception that they are cross-linked with melamines. Cure conditions for 1K clearcoats are 250°F for 30 minutes. Blocked isocyanate versions of 1K clearcoats also are available, but cure requirements necessitate 275°F or higher.

Significant research is being conducted on waterborne clearcoats in attempts to achieve better VOC emission levels. Waterborne clearcoat research is limited to water-dispersible acrylics because high gloss and good weatherability can be attained. Removal of water from the film, however, is a problem, as coalescence often begins prior to through-cure, often resulting in "solvent-popping." Much work remains to be performed on optical formulation parameters (i.e., solvent blend, anticratering surfactants, rheology control, etc.) in order to attain defect-free coatings.

Coatings for Plastics

Plastics use in automotive applications has expanded considerably over the past 20 years. In 1990, approximately 250 lb of plastic was utilized per automobile compared to < 100 lb five years earlier. With the increased use of plastics, coating research has concentrated on the painting of these often difficult substrates. Among the numerous prerequisites of coatings for plastic are: wetting of the substrate, uniform film development, a cure response below the heat distortion temperature of the plastic, adhesion to the substrate, inhibition

of substrate imperfections, and often a conductive nature to increase the transfer efficiency of subsequent topcoat application.

Plastic types are variable, ranging from flexible to rigid and from thermoplastic to thermoset. Because of these variations, the coating type applied to them also fluctuates, depending on cure response, solvent sensitivity, and modulus. The type of coating also depends upon the coated plastic end use. For example, topcoats can be applied directly to the plastic without the use of a conductive primer if two-toning or blackout areas are not part of the styling latitude. Adhesion promoters also are often utilized if a particularly difficult-to-adhere-to plastic is encountered (i.e., polypropylenes or thermoplastic olefins).

Conductivity in plastic coatings is achieved through the use of a conductive pigment, most often conductive carbon black. Conductivities on the order of 1 megohm/cm^2 are required for effective transfer efficiencies of subsequent topcoat applications. Without the use of a conductive primer, plastics in themselves would not be conductive, and hence would be very unreceptive to topcoat transfer efficiencies of greater than 20 percent.

The adhesion of a coating to plastic is related to the substrate wettability and also to physical–chemical forces. Both covalent (bond formation) and dispersive (London or van der Waals) forces are responsible for the adhesion of the coating due to physical–chemical forces.

Substrate wetting often is measured by contact angle measurements: the smaller the contact angle, the better the wetting of a coating to the plastic. The wettability of a coating on a plastic substrate is related to the substrate's surface tension: the lower the surface tension of the plastic, the more difficult it is to wet.

Polypropylene, in particular, has a very low surface tension (less than 32 dynes/cm^2). Because of the low surface tension, it is very difficult to wet out; and because of the apolar nature of the polypropylene backbone, it is a very difficult material to adhere to. Therefore, copolymers of ethylene propylene diene monomers (EPDM) with polypropylene have been developed. The higher the level of the EPDM in these thermoplastic olefins (TPO), the more paintable the substrate is.

Conventional coating approaches to TPO, which rely on the use of substrate pretreatments such as flaming, corona discharge, plasma, or chemical pretreatment, have led to a substantial oxidation of the surface which enhances adhesion. These pretreatments, however, have a limited service life before the surface reverts to its unoxidized, apolar condition.

Chlorinated polyolefins (CPOs) also are a mainstay of adhesion promotion to TPO. The CPO generally is applied as a dilute solution (5–35 wt. %) in aromatic solvents to the TPO to act as a "tie-coat," to which the topcoat is applied. These materials are often humidity-sensitive and only adhere to TPOs with EPDM levels of greater than 5 percent.

NEW CROSS-LINKING TECHNOLOGIES

The traditional cross-linking technologies utilized in the coatings market range from hydroxyl functional binders (acrylics or polyesters) cross-linked with melamines (1K) or isocyanates (2K) to drying oils that cure through oxidation of unsaturation in the backbone, to epoxy resins cross-linked with amino resins (amido amines, amines, etc.).

New cross-linking technologies, however, are being developed with hopes of achieving more environmentally acceptable coatings. Fluoropolymer coatings, epoxy-anhydride new curing technology (NCT), and siliconized-acrylics new enamel technology (NET), are a few of the new coating chemistries being introduced. Increasing demands by automotive producers have necessitated these advances. The need for properties such as chip resistance and resistance to "acid rain"[11] may result in further technological advances.

PRINTING INKS

Printing inks, utilized in printing and packaging applications, are colored, liquid materials consisting of essentially two components: the

colorant, an insoluble solid or a dye, and the vehicle, a liquid in which the colorant is suspended or dissolved. Recent years have seen significant progress in water-based inks, but compared with the phenomenal growth in water-based paints, the advances in water-based inks have had a much smaller impact on the industry. The upsurge of interest in water-based inks, which began in the mid-1970s, was stimulated by the need to conserve expensive petrochemical-derived organic solvents, attention to environmental pollution problems, and increasing pressures from health and safety legislation.

Inks can be applied to a wide variety of substrates, including plastics, paper, textiles, metals, and glass. In contrast to paint films, inks are applied very thin, varying from 2 to 30 microns depending on the print process. The print process consists of application of the ink to the substrate by a printing press of varying design. As such, inks are classified by the type of process through which they are applied. There are five classifications: flexographic, gravure, offset lithographic, letterpress, and screen.

The mechanisms of letterpress, offset lithographic/flexographic, and gravure printing processes are outlined in Fig. 27.3. In letterpress and offset lithographic processes, a thin film of ink is carried to the printing plate by means of a train of rotating rubber rollers. Both of these processes preclude the use of volatile organic solvents in the inks because of the swelling effect or attack on the rubber rollers and blankets and because of their high

Fig. 27.3. Mechanisms of letterpress, offset lithographic/flexographic, and gravure printing processes.[12]

evaporation rates, which could lead to premature drying or tack development of the inks on rollers, blanket cylinders, and printing plates. Water also is evaporated more rapidly than the high-boiling aliphatic-type solvents normally used in letterpress and offset lithography, and the evaporation is accelerated by the heat that can be developed (temperatures of 30°C or more) in the rollers after a period of running.

Several different types of drying mechanisms may be operative in printing of letterpress and offset lithographic inks (see Table 27.8). As shown in Table 27.8, the ink distribution systems of flexographic and gravure presses are quite different from those of letterpress or offset lithography. There are no long roller trains, and because the drying mechanisms depend mainly upon evaporation of volatile solvents components, it is not surprising that most of the development work has been on inks for these processes.

Inks that are printed by the flexographic and gravure printing processes are characterized by their extremely fluid nature, and are generally termed "liquid inks." When compared to the fluid nature of flexographic and gravure inks, lithographic and letterpress inks are much more viscous and pastelike.

In screen printing, the ink is forced through the open areas of a stencil supported on a mesh of synthetic fabric, stretched across a frame. The ink is forced through the mesh onto the substrate underneath by drawing a squeegee across the stencil. The inks have to remain fluid and coherent on the mesh before printing, transfer through the mesh without its sticking to the substrate, flow to make the ink film continuous and so avoid the pattern of the mesh but maintain the image, and then dry rapidly once applied to the substrate. Inks that are able to print by this method are of household paint consistency; they are intermediate in viscosity between the fluid flexographic and gravure inks and the pastelike lithographic inks.

An ink must be transformed to a solid from its liquid supply form once on the substrate. This change of state is referred to as an ink drying, and can be physical (absorption or evaporation) or chemical (oxidative or photochemical radiation–induced). There are several forms of radiation used to dry ink, including ultraviolet, infrared, electron beam, microwave, and radio-frequency.

PIGMENTS

Pigments are generally incorporated into paints to affect properties associated with appearance such as color, opacity, gloss, metallic look, and depth. In addition, pigments commonly are used to protect the substrate against corrosion, attack by microbes, or barnacles and to retard flammability. Other desirable properties can be obtained by the addition of pigments. For example, pigments can be used in paints to control flow and leveling as well as either spray viscosity or brushability, depending upon the type of system. In addition, pigments are added to paint to control the cost of raw materials.

One of the most important properties of pigments, however, is the ability to cover or "hide" the surface being painted. This property is commonly referred to as "hiding power" or opacity. The fewer coats required to hide the surface, the higher the hiding power or the pigment's opacity. Pigments of different colors have different degrees of opacity. In general, darker colors have a higher hiding power than lighter colors, in part because of their relative opacities. In practice, this trend is also dependent upon the refractive index of both the pigment and the binder; that is, the difference in refractive index between the pigment and the binder determines the hiding power of the paint. Hence, the larger the refractive index difference, the higher the hiding power. In addition, the particle size of the pigment critically affects the hiding power of the pigment. There is an optimum particle size range, 0.2 to 0.4 micron (which is approximately one-half the wavelength of light in air), which affords maximum light scattering and hiding.[13]

Pigments ideally should be nonsoluble solids with low oil absorption to facilitate dispersion and allow maximum pigment

loading, which would also reduce the VOC content of the paint. In addition, pigments should be chemically inert, thermally stable, pure in their composition, and nontoxic.

Pigments can be classified into two main categories: inorganic pigments and organic pigments. Among the more common inorganic pigments are white hiding pigments, extender pigments or fillers, blacks, colors, and metallic pigments. Organic pigments consist of colors and the carbon blacks.

Inorganic Pigments

Inorganic pigments generally hide better than organic pigments when dispersed in paint. Inorganic pigments impart flowability to paint because they do not tend to swell in the presence of common paint solvents. In addition, inorganic pigments do not bleed in organic solvents, and they afford excellent color permanence, especially in exterior paints. Inorganic pigments are thermally stable and generally impart heat resistance to the coating.

White pigments are important not only as the principal pigment in white paint but also as hiding pigments in many shades and tinting colors. The most commonly used of the white pigments are the nonreactive titanium dioxides and the reactive zinc oxide. These pigments are designated reactive or nonreactive, depending on whether or not they interact chemically with the binder or vehicle portion of the paint.

Titanium dioxide is available in two crystalline forms—anatase and rutile. The anatase form has about 75 percent of the hiding power of the rutile form and is less resistant to chalking than the rutile form. Because of its high refractive index (2.76), titanium dioxide imparts excellent opacity to the paint. Because the hiding power of the titanium dioxide is so high, good opacity can be achieved by the addition of small amounts of the pigment to the paint, thus not adding significantly to the cost of the paint. The rutile form is available in grades designed specifically for enamels, flats, and solvent-based and water-based coatings.[14]

Although zinc oxide has about 15 percent of the hiding power of titanium dioxide and has a fairly high cost, it contributes a number of important properties to a paint. Zinc oxide can form zinc soaps in some formulations, which can aid in pigment wetting and grinding. It also can be utilized to improve film hardness and color retention of the film on aging. The most common use of zinc oxide is in exterior finishes to reduce chalking and the growth of mildew in house paints.[14]

Extender pigments, because of their relatively low refractive index, exhibit poor hiding power and must be used in conjunction with pigments of high hiding power. The principal function of extender pigments in paint is to reduce the raw material cost. They also function to reduce gloss, as well as to control texture and viscosity. Calcium carbonate, more commonly known as whiting pigment, is available in a variety of particle sizes and surface treatments and is used to control sheen, degree of flow, degree of flatting, and tint retention.[14] Magnesium silicate or talc generally is used because of its low cost and excellent durability. Talc imparts good anti-settling properties and low sheen. Aluminum silicate (China clay) is used mainly in water-based systems. China clay is easily dispersed with a high-speed disperser and may even add to the hiding power of water-reducible paints. Other extenders such as diatomaceous silica, regular silica, barytes, and mica also are frequently used to impart specific properties to coatings, which are determined mainly by the particle size and shape of the pigment.[13,14]

Most of the black pigment used in coatings is organic carbon black. There are, however, a few inorganic black pigments such as black iron oxide, which is used as a tinting black having brownish tones and also is used in primers. In addition, some inorganic black pigments such as mineral or thermal blacks are used as low-cost black extender pigments.[14]

Iron oxide pigments are low-cost, high-opacity pigments that impart excellent properties to coatings. These pigments are easy to disperse in both water and oil, are alkali- and chemical-resistant, and hence can be utilized

in both water-based and solvent-based systems. Although both the red and yellow iron oxides afford rather dull colors, they possess high tinting strength and are popular shading colors. In addition, iron oxide pigments are extremely durable, colorfast, and nonbleeding.

Lead chromate pigments are mainly used as corrosion inhibitive pigments. These pigments are of low opacity and can be combined with higher-opacity pigments for use in primers. Zinc chromate, also an anticorrosion pigment, is used in metal pretreatments such as "wash primer," which also promotes adhesion to aluminum and steel.

Iron blue, one of the most widely used pigments, combines low cost, good hiding power, and high tinting strength. This blue pigment exhibits good exterior durability and good heat resistance. Iron blue should not be used in water-based paints because it has very poor alkali resistance.

Ultramarine blue has poor opacity but high tinting strength, high heat resistance, and good alkali resistance. It is primarily used as a tinting color, and when added to whites makes then look whiter by giving them a bluish cast.[14]

Chrome greens range in color from light yellow greens to dark blue greens. They are bright, durable, and economical and have high hiding power. They are used less frequently of late because of their lead content. Chrome greens are not suggested for use in aqueous systems because of their poor alkali resistance.

Chrome oxide green is a dull pigment having excellent exterior durability and very good chemical as well as alkali resistance. Thus, this pigment is suitable for use in either water-based or solvent-based systems. Chrome oxide green has moderate hiding power and is easy to emulsify, and its infrared reflectance properties make it valuable for use in camouflage paints.[14]

Organic Pigments

Organic pigments are insoluble solids, and thus are differentiated from dyes, which are colored organic compounds that are soluble in the media in which they are used. Pigments require a binder in order to be used as colorants. Organic pigments that are insoluble in their pure form are referred to as "toners." Organic pigments that require precipitation by an inorganic base such as a metal to be insoluble are referred to as "lakes."

Organic pigments have greater decorative value than inorganic pigments; that is, organic pigments have a greater aesthetic appeal due to higher chroma. The higher tinting strength of organic pigments partially offsets their higher cost. Organic pigments have a smaller particle size than inorganic pigments, and a higher oil absorption. In addition, organic pigments are generally less toxic than inorganic pigments.

Hansa yellows are insoluble azo pigments having a wide range of shades. They are generally excellent in masstone lightfastness although not good enough in tint lightfastness to warrant outdoor use. Compared with the chrome yellows, they have less hiding power and greater tint strength. The Hansa yellows are most commonly used in water-based coatings and air-dried coatings. Benzidine yellows are no longer used because they have been shown to be carcinogenic. Nickel azo yellows are transparent pigments used in greenish yellow enamels and tinting enamels. These pigments have better permanency than the benzidine yellows although they tend toward darker transparent masstones. Nickel azo yellows impart excellent alkali resistance and fairly good bleed resistance. The tinting strength of these pigments increases upon aging in the formulation. Vat yellows are comprised of flavanthrones and anthrapyrimidines. Of these, the flavanthrones are the more red in hue. The vat yellows have very good lightfastness in weak tints, and they impart very good heat resistance and bleed resistance to the coatings. Tetrachloroisoindolinone-based pigments, although relatively expensive, are of very high quality, imparting good lightfastness to the coatings.

Benzidine orange pigments are bright, strong pigments that exhibit good alkali, heat, and bleed resistance. There are both opaque and transparent versions of these pigments,

and they are fairly lightfast although they are not recommended for exterior use. Benzidine orange is no longer used because of its potential carcinogenic nature. Dianisidine orange pigments are more bleed-resistant than the benzidine oranges but less colorfast. Dinitroaniline orange pigments have better permanancy than the two orange pigments described above. Although the dinitroaniline orange pigments bleed in most solvents, they are very stable in aqueous systems and exhibit superior lightfastness. Tetrachloroisoindolinone, brominated anthanthrone, and perinone oranges are expensive but very high-quality orange pigments characterized by their excellent masstone and exterior lightfastness.

Copper phthalocyanine greens afford very clean, semitransparent masstone, nonbleeding tints ranging from bluish green to yellowish green. These pigments have excellent heat and chemical resistance, are very lightfast in both masstone and tint, and have high tint strength. Pigment green B is an iron salt of nitrosobetanaphthol that affords a semitransparent masstone having a dull yellowish tint. Although this material is not lightfast enough for exterior applications, it does have excellent alkali resistance and is suitable for use in aqueous pigment dispersions.

Phthalocyanine blues are very clean, semitransparent, nonbleeding pigments yielding high tint strength. As with the phthalocyanine greens, the blues have excellent heat and chemical resistance and are very lightfast in both masstone and tint. Also as with the greens, these blue pigments at high concentrations may bronze upon exterior exposure although this property is vehicle-dependent. These characteristics apply to the crystallization and flocculation grades. These pigments are highly recommended for use in water-based systems. Indanthrone blues are expensive vat-type, nonbleeding, semitransparent masstone pigments having excellent masstone and tint lightfastness. Although they exhibit very good heat and chemical resistance, these pigments find limited use in coatings.

Carbazole dioxazine violets are semitransparent masstone pigments having high tint strength, excellent heat and bleed resistance, and good lightfastness. Isoviolanthrone violets afford clean tints exhibiting excellent chemical resistance and very good lightfastness, but they bleed badly in paint solvents.

Toluidine reds are insoluble azo pigments ranging from light to deep reds having excellent chroma and masstone lightfastness. Although these pigments hide very well and exhibit very good alkali resistance, they are not suitable for exterior use and bleed in paint solvents. Toluidine reds have been replaced to a large extent by a combination of molybdate orange and quinacridone or rubine toners. This combination affords better gloss and gloss retention, better bleed resistance, and less haze. Compared with the toluidine reds, this combination does sacrifice some cleanliness and brilliance of tint as well as some masstone lightfastness. Para reds are insoluble azo pigments that are deeper, less bright, and not as masstone-lightfast as the toluidine reds. They have good opacity and good alkali resistance, but they bleed very badly in paint solvents. These pigments can be utilized in inexpensive air dried water-based systems. Quinacridones possess hues from gold to deep maroons and violets. These reds afford excellent bleed and heat and alkali resistance, as well as very good tint and masstone lightfastness.

Carbon black is probably the most widely used pigment in the coatings industry. It is popular both as a straight shade and as a tinting color. It comes in a wide variety of particle sizes and surface treatments, the choice depending upon required jetness, tone, ease of dispersion, and of course cost. Carbon black is characterized by excellent hiding power or high opacity, high tint strength, excellent durability, and very good lightfastness. Large-particle-size carbon blacks are chosen for their ease of dispersion and low resin demand, as well as for being less susceptible to floating and flooding than the finer particle sizes. They should, however, be used in applications that do not demand high levels of jetness. The finer-particle-size, higher-surface-area carbon blacks afford extremely jet, very durable, high gloss coatings. These

carbon blacks would be very suitable in automotive topcoat applications. The fine-particle-size carbon blacks also are used frequently in conductive coatings. Although all carbon blacks are good conductors, differences exist based on surface area, structure, and volatile content. In general, however, conductivity of the carbon black increases as the surface area and structure increase and as the volatile content of the carbon black decreases. The structure of the carbon black is defined as the agglomeration of the fine particles into treelike or "dendritic" shapes.

Pearlescent Pigments

Pearlescent pigments are comprised mainly of mica platelets coated with a thin layer of titanium dioxide. Coprecipitating titanium dioxide and other metal oxides onto the mica results in pigments exhibiting a colored luster effect. These pigments have a very high refractive index, indicating high hiding power or high opacity. In addition, because of their light-reflective properties they impart some exterior durability to the coating; and because of their small, thin, flat structure and high refractive index they give the coated object a three-dimensional effect, that is, the appearance of depth. This is quite different from the apparent surface gloss due to the use of metallic pigments. Pearlescent pigments are available in a range of particle sizes and characteristics, and provide a variety of effects. These include a satin-luster appearance resulting from fine-particle-size pigments. These pigments are alkali- and acid-resistant and extremely durable when coated with rutile titanium dioxide. Pearlescent pigments are readily dispersible in water and afford stable metal-like water-based coatings.

Aluminum Pigments

Aluminum pigments generally are provided in paste form to eliminate the hazards of dust explosion or ignition during handling and storage, as occurs with the aluminum powders. In addition, an aluminum paste is very economical because of its high hiding power, excellent durability, and low specific density. Aluminum paint films provide excellent protection of substrates because the aluminum flakes align in many parallel layers. The layers of aluminum flake are staggered, thus yielding a tortuous path for the transmission of gases, chemicals, light, and moisture. Therefore, it takes longer for these materials to penetrate the film to get to the substrate. Aluminum paints made up of the leafing-type of aluminum paste reflect 75 to 80 percent of light and radiant heat, and are commonly used as exterior maintenance-type coatings or decorative coatings. In fact, the minute flakes in aluminum paste impart a smooth, brilliant appearance to the film. The nonleafing type of aluminum flake can be tinted with other various colored pigments to effect brilliant finishes such as those used in automotive metallic topcoats.

LACQUERS

Some time before 200 B.C., the Chinese used an exudate from the conifer *Rhus verniciflua* (later known as the varnish tree) as a coating. The process was further developed in Japan, where by A.D. 400 the product was known as Japanese lacquer. These lacquers are a type of oleoresin that dries by oxidation in a damp atmosphere.[15]

Lacquers are defined as nonconvertible coatings; that is, lacquer film formation does not depend upon any chemical reaction. Film formation is dependent only upon the evaporation of solvents, and the formed film remains soluble in the parent solvent blend. Lacquers generally are spray-applied, taking advantage of the solvents' ability to dissolve the previously applied coat. Chlorinated rubbers and nitrocellulose esters are typical examples of lacquers. Nitro-cellulose lacquers can form glossy coatings having excellent adhesion when other suitable resins are added that are compatible in all proportions. Two types of cellulose nitrates are available, a low-viscosity type that affords high-solids, high-film-build coatings and a high-viscosity, low-solids type. The primary use of pigmented nitro-cellulose

lacquers has been in the car refinish and the wooden domestic furniture markets. The chlorinated rubber lacquers are used principally as protective coatings for steel and other vulnerable materials in particularly harsh environments.

Vinyl lacquers are used mainly where a high degree of chemical resistance is required; these lacquers are based on vinyl chlorides and vinyl acetates. Acrylic lacquers are based on methyl methacrylate and methyl acrylate polymers and copolymers. Other esters of acrylic and methacrylic acid also may be used to make nonconvertible film formers. Judicious selection of these acrylic acid or methacrylic acid esters allows one to produce film formers with specifically designed properties such as hardness, flexibility, gloss, durability, and heat and chemical resistance. Acrylic lacquers, however, are not noted for their water resistance. The principal uses of acrylic-type lacquers are fluorescent and metallic paints, car refinish applications, clear lacquers and sealers for metals, and protective coatings for aircraft components and for vacuum-deposited metals, as well as uses in pigmented coatings for cabinets and appliances.

Vinyl emulsion paints that contain a nonconvertible resin dispersed in water along with pigment can be considered as waterborne lacquers if film formation results from evaporation of the water followed by coalescence of the binder in which the pigment is dispersed. As environmental regulations become more strict, the role of traditional lacquers will essentially go the way of the dinosaur. The only allowable lacquers or nonconvertible coatings eventually will be of the water-based type. Indeed, as environmental regulations become more restrictive, the challenges facing the coatings chemist and the formulator become ever more demanding.

REFERENCES

1. *Chemical Week*, pp. 50–73 (Oct. 31, 1990).
2. *Chemical Week*, p. 30 (Oct. 18, 1989).
3. Sawyer, R., *Paint and Coatings Industry*, p. 25 (Sept. 1990).
4. *Industrial Finishing*, p. 31 (Mar. 1991).
5. Schulte, S., *Products Finishing*, p. 28 (June 1990).
6. *Paint and Coatings Industry*, p. 25 (Apr. 1988).
7. *Chemical and Engineering News*, p. 54 (Sept. 17, 1990).
8. Bocchi, G., *Modern Paint and Coatings*, p. 36 (Nov. 1989).
9. Besold, R., Reisser, W., and Roth, E., Waterborne and Higher Solids Symposium, New Orleans, LA, Feb. 6–8, 1991.
10. Akkerman, J. M., *Prog. Org. Coatings*, **17**, 53–68 (1989).
11. Schrantz, J., *Industrial Finishing*, p. 14 (July 1990).
12. Hutchinson, G. H., *JOCCA*, **68** (12), 307 (1985).
13. Martens, C. R., *Riegel's Handbook of Industrial Chemistry*, J. Kent (Ed.), p. 787, Van Nostrand Reinhold, New York, 1983.
14. Singer, E., *Paint Handbook*, pp. 3–23 to 3–32, McGraw-Hill, New York, 1981.
15. Skolnik, H., *Organic Coatings: Their Origin and Development*, B. Seymour and H. F. Mark (Eds.), p. 16, Elsevier, New York, 1990.

28

Synthetic Nitrogen Products

Lawrence D. Pesce* and William R. Jenks**

Nitrogen products are among the most important chemicals produced in the world today. The largest quantities are used as fertilizers, but nitrogen products also find very important uses in the manufacture of nylon and acrylic fibers, methacrylate and other plastics, foamed insulations and plastics, metal plating, and gold mining, to list a few. There are virtually no nitrogen compounds that can be mined or distilled from crude oil or any other natural source. Therefore, economical fixation of nitrogen from the atmosphere has been a never ending quest. It has been a difficult task because elemental nitrogen is comparatively unreactive. It combines with few other elements, and then usually only under drastic conditions of high pressure and/or high temperature.

NITROGEN FIXATION

Most nitrogen products have been based on nitrogen oxides and ammonia; so the history of fixation of nitrogen as nitrogen oxides and ammonia is of interest.

Nitrogen Oxides

In 1780, Henry Cavendish, an English scientist, used an electric spark to combine atmospheric nitrogen and oxygen to make nitrogen oxides. In 1902, Bradley and Lovejoy performed experiments at Niagara Falls aimed at developing a commercial process using an electric arc to produce nitrogen oxides. The experiments failed, apparently because the electric arc "flame" area was too small, and product gases were not removed from the reaction zone fast enough. Later, with the help of these results, Birkeland and Eyde, in Nottoden, Norway, developed the first large-scale manufacture of nitrogen oxides from air using an arc spread out to form a flame.

Nitrogen oxides also can be produced from the nitrogen and oxygen in air by raising the air to very high temperatures. Much experimental work has been done to develop a commercial process, but to date the concentration of the nitrogen oxides has been too low for economic success. Also, the high

*Senior Engineering Associate, E. I. du Pont de Nemours & Co.
**Retired, E. I. du Pont de Nemours & Co.

temperatures required make commercial application of this process unattractive. A typical high temperature process operates at 2000°C; therefore, the choice and the maintenance of materials of construction are critical. Feed to the process is raised to the reaction temperature by means of a batch pebble preheater. The pebbles first are heated by burning a fuel such as natural gas or fuel oil, and then the air is passed through the preheater. Pebble attrition is another serious problem in this process.

Other processes have been proposed for nitrogen fixation, but none of these has approached a commercially attractive stage. Recently, for example, it was shown that low energy ultraviolet photons can fix nitrogen at room temperature under oxidizing conditions on the surface of a titanium dioxide–containing material.[1] Nitric oxide is formed.

Research continues on the formation of nitrogen oxides. The electric arc developed in Norway reportedly is still in use because of the inexpensive supply of electricity; but today the high temperature fixation of nitrogen as oxides remains an environmental problem rather than a commercially attractive process, as large quantities of polluting nitrogen oxides are produced by fixation in high temperature combustion processes such as power plants, automobile engines, and home furnaces.

Ammonia

Early in the twentieth century the manufacture of synthetic ammonia first succeeded with the development of the Haber process. In this process a mixture of nitrogen and hydrogen is passed over a contact catalyst at a moderately high temperature and under pressure and a portion of the elemental gases is converted to ammonia. Later modifications to this process have been so successful that the Haber process now is more important than all other synthetic ammonia processes combined.

Ammonia and some of its salts and derivatives are valuable fertilizers, for which there is a continuing need. Ammonia also is a very important building block for many other chemical commodities.

As noted earlier, much effort has been devoted over the years to developing methods for nitrogen fixation. These methods have included the use of dicyclopentadienyltitanium dichloride in the presence of ethyl magnesium bromide to use enzymes' ability to fix nitrogen at room temperature and atmospheric pressure. Another unique method[1] for fixation as ammonia involved feeding purified nitrogen and hydrogen into an electrolyte solution of an alkali. A catalyst of powdered nickel pellets promoted with cobalt is used in the process. It was reported that conversions of 4.2 percent from the gases per pass were achieved at a reactor temperature of 50 to 80°C.

Nitrogen also can be fixed by being passed over metals at suitable temperatures to form nitrides, which when treated with steam will yield ammonia. The best known process using this chemistry is the Serpek process, introduced in France for the manufacture of aluminum nitride.

Today the Haber process remains the most commercially attractive ammonia process although it has high compression costs, and a large expenditure of energy is required to produce the feed hydrogen.

Other Processes

Atmospheric nitrogen also can be fixed by the calcium cyanamide process, which is based on the fact that metallic carbides, particularly calcium carbide, readily react with nitrogen to form the metal cyanamide. Fairly pure nitrogen is reacted with calcium carbide to form the cyanamide:

$$CaC_2 + N_2 \xrightarrow{1000°C} \underset{\text{Calcium cyanamide}}{CaNCN} + C$$

The process was developed in Germany by Frank and Caro in the mid-1890s and currently is in use in many countries including Canada and the United States. Calcium cyanide can be produced by a similar reaction with the addition of sodium chloride as a

"catalyst":

$$CaC_2 + N_2 \xrightarrow{NaCl} \underset{\text{Calcium cyanide}}{Ca(CN)_2}$$

The principal use for calcium cyanide is in mining gold.

Nitrogen can be fixed as metal cyanide by being passed into a reactor containing an alkali and coal. The products are cyanides and ferrocyanides. This process has no commercial importance at present.

Nitrogen will react with acid azides in the presence of rhodium or iridium catalysts. Also, organic complex compounds have been shown to fix nitrogen in the presence of enzymes. At present, these also have no commercial value.

NITROGEN CONSUMPTION

During the period 1955 to 1965, world consumption of nitrogen as a fertilizer increased from 8 million tons to 16.3 million tons. From 1965 to 1975, consumption increased to 39 million tons. By 1990 the consumption was 102 million tons.[2,3]

World capacity for the production of nitrogen as a fertilizer has increased from 1975 to the present, and is expected to continue to grow because of new installations in the former Soviet Union, Asia, and Eastern Europe. Up until 1972, nitrogen fertilizer consumption in North America exceeded that in Asia. However, by 1985, consumption in Asia was 1.4 times that in North America and 1.6 times that in the Soviet Union.

The per capita consumption of nitrogen in the United States is about 80 lb/year, whereas it is only 50 lb/year in Europe and 8 lb/year in China. With an exploding population in the Far East, a total of 30 or more plants will be needed if the per capita consumption approaches that in the United States. Hence, there is potentially a great need for new construction and engineering services to meet the expansion of the industry.

Nitrogen capacities in various parts of the world for 1975 and 1990 are given in Table 28.1. Materials consumed as nitrogen fertilizer in the United States are given in Table 28.2, for the years 1965, 1970, 1977, and 1990. Ammonia consumption, both by direct application and in solution, shows some decrease. Urea continues to be of interest because of its high nitrogen content (i.e., 46.7% N vs. 35% N for ammonium nitrate). Consumption of ammonium sulfate continues to decrease because it is a by-product and contains only 21.2 percent nitrogen.

TABLE 28.1 Capacities for Fertilizer Nitrogen[2,3]

Area	*1000 M Tons N* 1975	1990
Asia	15,093	28,900
North America	14,850	17,000
Western Europe	14,360	16,000
U.S.S.R. and Eastern Europe	18,749	31,500
Latin America	2,672	5,100
Africa	832	2,700
Oceania	446	550
Total	67,002	101,750

TABLE 28.2 Selected Nitrogen Fertilizer Materials (Millions of Short Tons)[2,3]

Material	*1965*	*1970*	*1977*	*1990*
Ammonia	8.9	13.8	17.6	15.0
Ammonium sulfate	0.71	0.60	0.47	—
Ammonium nitrate	4.1	5.4	5.7	4.6
Urea	1.0	2.6	4.2	4.4

AMMONIA

With no economical nitrogen fixation via nitrogen oxides having been discovered, ammonia has developed into the most important, and practically the only, building block for all synthetic nitrogen products worldwide (Fig. 28.1). Worldwide, the production of ammonia exceeds 265 billion lb. About 85 percent of this production is used for fertilizers, with the remainder going to all other uses.

Of current world ammonia capacity—about 140 million short tons—70 percent is based

Fig. 28.1. Synthetic nitrogen products.

on natural gas, 6 percent on plants partially converted from oil to natural gas, 10 percent on naphtha and fuel oils, 4 percent on by-product gases from coke ovens or refinery gas, and 10 percent on coal. In the future, coal will become of increasing importance as natural gas supplies are consumed, and natural gas prices escalate.

Since 1963, there has been a revolution in ammonia-manufacturing technology. The advent of large single-train plants has resulted in a large increase in production capacity, the

shutdown of a number of smaller plants, and a reduction in manufacturing costs. The total capacity of U.S. ammonia producers in 1990 was estimated at 15,490 thousand metric tons/year.[4]

Prior to World War II, ammonia production capacity remained relatively stable. During the war the need for explosives caused an increase in the production of ammonia for nitric acid manufacture. Then, after the war, available ammonia capacity was directed to the manufacture of fertilizers. Accordingly, there was a rapid increase in fertilizer consumption. Therefore, the advantages of fertilizer were emphasized, and production capacity increased by leaps and bounds.

From 1940 to 1950 the number of ammonia plants doubled; then from 1950 to 1960 the number more than doubled again. It was in the early 1960s that the single-train concept was developed, and many smaller plants became uneconomical and were shut down. Capacity tripled in the period from about 1958 to 1968. U.S. production declined in the 1970s, and although it has leveled off at near 16 million metric tons/year, world demand continues to rise. Worldwide production, including the U.S., increased 70 percent from 1976 to 1990 to 121 million metric tons/year.

All ammonia manufacturing processes are based on the synthesis of ammonia from nitrogen and hydrogen. Thus the many processes are differentiated, among other factors, by the method of producing the hydrogen and nitrogen for the synthesis.

Preparation of nitrogen can be by separation from air in a standard "air" plant, or air can be fed to the ammonia unit where the oxygen is used to combust a portion of the hydrocarbon feed, leaving the nitrogen for ammonia synthesis. The latter is the dominant type.

With the advent of "platforming" of petroleum products to give unsaturated compounds for high test gasoline, large volumes of hydrogen off-gas have been produced.[5] The off-gas from "platforming" operations frequently contains over 90 percent hydrogen. In such cases, the processing requirements for ammonia manufacture are greatly reduced. It is only necessary to wash out the higher hydrocarbons and scrub the gas with liquid nitrogen to remove methane and carbon monoxide for the hydrogen–nitrogen mixture to be ready for compression and ammonia synthesis. Currently, most "platforming" hydrogen is used in hydrocracking units, and many petroleum producers are actually short of hydrogen. Therefore, most hydrogen for ammonia manufacture is prepared from hydrocarbon feeds.

Some physical properties of ammonia and densities of aqueous solutions of ammonia are given in Tables 28.3 and 28.4, respectively.

TABLE 28.3 Physical Properties of Ammonia

Property	*Value*
Molecular weight	17.03
Boiling point, °C	−33.35
Freezing point, °C	−77.7
Critical temp., °C	133.0
Critical pressure, kPa (psi)	11,425 (1657)
Specific heat, J/(kg·K)	
0°C	2097.2
100°C	2226.2
200°C	2105.6
Heat of formation of gas, ΔH, kJ/kg mol, (Btu/lb·mol)	
0°K	−39,222 (−16,862)
298°K	−46,222 (−19,872)
Solubility in water, wt. %	
0°C	42.8
20°C	33.1
40°C	23.4
60°C	14.1
Specific gravity[a]	
−40°C	0.690
0°C	0.639
40°C	0.580

[a]Anhydrous ammonia.

TABLE 28.4 Densities of Aqueous Ammonia at 15°C

Ammonia, wt. %	*Density, g/liter*
8	0.970
16	0.947
32	0.889
50	0.832
75	0.733
100	0.618

HYDROGEN PRODUCTION

Preparation of hydrogen is done by two general methods. The hydrocarbon feed can be reformed with steam to produce synthesis gas, which is a mixture of carbon oxides and hydrogen, or the hydrocarbons can be partially oxidized to produce the synthesis gas. Each process has advantages that are primarily dependent on the hydrocarbon feed to be used. The hydrocarbon feed can be a range of materials, from methane in natural gas through all the higher hydrocarbon gases such as propanes and butanes to fuel oils and naphtha to coal.

In the United States, natural gas continues to be the preferred source of hydrogen and energy. Coal is a good source of energy, but a poor source of hydrogen. When coal is used, the hydrogen is obtained from water. Nuclear fission could supply abundant energy, but the source of hydrogen for a fission process is yet to be determined.

With natural gas as the raw material, plants depend on steam reforming to produce hydrogen and carbon monoxide. Air is bled into a secondary reformer to supply nitrogen. The carbon monoxide is then shift-reacted with steam to produce hydrogen and carbon dioxide. The carbon dioxide is removed, and the last traces of carbon monoxide and carbon dioxide are converted to methane, which is not a poison in the ammonia synthesis step. The purified hydrogen and nitrogen are compressed and fed to the ammonia synthesis loop.

Partial combustion of natural gas and naphtha with oxygen is used to produce hydrogen and carbon monoxide.[6] Processes based on this initial step convert the carbon monoxide with steam to carbon dioxide and hydrogen directly after partial combustion. Most of these installations are in Europe and India.

The carbon dioxide is removed; the gas is dried and scrubbed with liquid nitrogen to remove residual carbon monoxide, argon, and methane; and then it is compressed and sent to ammonia synthesis. The removal of argon and methane improves the utilization of hydrogen in the synthesis loop.

Partial oxidation has the advantage that several different feedstocks such as natural gas, naphtha, fuel oil, and petroleum fractions can be used. In addition, the H_2/CO ratios of the synthesis gas can be made more favorable for end products such as methanol, oxo alcohols, and acetic acid (pure CO), if these alternatives are desired.

Figure 28.2 illustrates a process that starts with coal or oil as a raw material. The process includes coal handling, preparation, and pulverization and partial oxidation of coal or oil to synthesis gas, followed by carbon and ash removal, carbon monoxide conversion, carbon dioxide and H_2S removal, low temperature scrubbing with liquid nitrogen, and compression for ammonia synthesis. Air separation and sulfur recovery are steps unique to coal or heavy oil feedstocks.

For an ammonia plant at a new location where by-product hydrogen is not available,

Fig. 28.2. Ammonia from coal or oil.

TABLE 28.5 Synthesis Gas H_2/CO Ratio

Method of Manufacture	*H_2/CO Ratio*
Oxygen-coke-steam	0.6
Air-coke-steam	0.9
Oxygen-coal-steam	1.0
Oxygen-fuel oil-steam	1.0
Propane-steam	1.33
Methane-oxygen	1.7
Methane-oxygen-steam	2.3
Methane-steam	3.0–5.0

selection of the most economical process will depend to a large extent on the raw material used. Recent innovations[7,8] have been made to improve the energy efficiency of the natural gas fired plants.

A list of processes for producing hydrogen is given in Table 28.5, along with the hydrogen-to-carbon monoxide ratio of the synthesis gas produced. When synthesis gas is made by the gasification of coke with steam and oxygen, the hydrogen-to-carbon monoxide ratio is 0.6; with the intermittent gasification of coke with air and steam, the synthesis gas has a ratio of 0.9. By using excess steam with the reforming of methane, the hydrogen-to-carbon monoxide ratio can be increased to as high as 5. Thus by selection of the proper process, equipment requirements for conversion of carbon monoxide to hydrogen with steam can be reduced.

Very often, the raw material that is available, such as natural gas, coal or oil, will dictate the type of equipment or process employed. Table 28.6 gives raw material requirements for producing synthesis gases by various processes.

Selection of a process for hydrogen manufacture depends not only on the raw material and its cost but also on the scale of operation, the purity of the synthesis gas produced, the pressure level of the natural gas, and other outside consumption for carbon monoxide and hydrogen.

TABLE 28.6 Synthesis Gas Raw Material Requirements

Method of Manufacture	*Feed per MCF of ($CO + H_2$)*	
Air-coke-steam	40 lb	coke
	80 lb	steam
	2200 cu ft	air
Oxygen-coke-steam	30 lb	coke
	300 cu ft	oxygen
	34 lb	steam
Coal partial combustion	37 lb	coal
	350 cu ft	oxygen
	37 lb	steam
Natural gas partial combustion	360 cu ft	natural gas
	260 cu ft	oxygen
Natural gas reforming	285 cu ft	natural gas
	230 cu ft	natural gas for fuel
	60 lb	steam
Propane reforming	3 gal	propane
	1.67 gal	distillate for fuel
	380 lb	steam[a]
Naphtha reforming	14.7 lb	naphtha
	5.8 lb	naphtha for fuel
	105 lb	steam
Fuel oil partial combustion	20.5 lb	oil
	320 cu ft	oxygen
Electrolysis	163 kWh	
Steam-iron	290 lb	steam
	0.3 lb	ore
	206 cu ft	fuel gas
	65 lb	coke
Coke-oven gas	353 lb	coal[b]
"Platforming" off-gas	84 gal	feedstock

[a] Includes CO shift to CO_2 and H_2 with steam.
[b] Excludes $CH_4 + CO$ in coke-oven gas which can be converted to H_2O.

Feedstock Purification

Because most of the ammonia plants today are based on the reforming of natural gas with steam over a nickel catalyst to produce hydrogen, it is imperative that the sulfur content of the feedstock be reduced to as low a level as possible. Natural gas for transportation by pipeline usually is desulfurized to minimize pipeline corrosion. Hence, by the time it reaches an ammonia producer, the only sulfur present is that used to odorize the gas. On the other hand, naphtha or fuel oil that is used as a raw material may require drastic treatment to remove sulfur.

For natural gas purification, the usual treatment is to pass the gas over specially prepared activated carbon to remove the sulfur. The treatment takes place at ambient temperature and at the pressure level used in the reformer or natural gas line. The higher

the pressure, the more effectively the carbon is used.

The activated carbon may be impregnated with iron or copper to enhance the removal of sulfur from the gas. About 100,000 to 200,000 cu ft of natural gas/cu ft of carbon can be treated. Then the carbon must be regenerated with high temperature steam at low pressure. The temperature of regeneration must exceed the boiling point of the adsorbed compound with the highest boiling point. Usually a temperature of 250 to 300°C is satisfactory. A two-bed system normally is used in the activated carbon process; one is in operation while the other is being regenerated.

For desulfurization of naphtha, a more complicated process is required. Hydrodesulfurization quite often is used. About 0.5 mole of H_2 is mixed with 1 mole of vaporized naphtha or 250 scf (Standard Cubic Feet) per barrel, depending upon the sulfur and olefin content. The mixture is preheated to 320°C and passed over a cobalt–molybdenum catalyst, where the olefins are hydrogenated to paraffin hydrocarbons and the sulfur compounds are reduced to H_2S. The gas then is passed over a sulfur adsorbent such as iron or zinc oxide. It may or may not be necessary to condense the naphtha, depending upon the amount of hydrogen used and the need to remove it from the naphtha.

The hydrogenation of olefins to paraffin hydrocarbons is advantageous because olefins otherwise would readily crack to carbon in the next step (reforming) of the process. U.S. Patent 3,106,457[9] gives conditions under which olefins will crack to carbon. The minimum steam-to-carbon ratio reported to avoid carbon is given as:

$$(Ri)_0 = (0.244p + 8.15 \times 10^{0.0052M})$$

where $(Ri)_0$ is the minimum steam to carbon ratio (moles per mole organic carbon atom), p is the operating pressure in psia, and M is the molecular weight of the olefin.

For preparation of coal, after tramp iron and wood are removed, it must be crushed, sized, pulverized, and in some cases slurried with water for injection into the reactors. It also may be profitable to use some applicable method for the type of coal available to remove sulfur-bearing gangue.

Coal selection is important for proper gasifier operation. For slagging, liquid-bottom gasifiers, the melting point of the ash will be critical because a high melting point may result in a frozen slag-tap opening. In other cases, with fixed-bed gasifiers, caking coals may result in agglomeration of ash and coal, producing blow holes that will allow oxygen and steam to break through into the product gas. A high ash content also will result in poor yields and increased costs for ash handling. High sulfur coals may overload the gas cleanup section of the plant if it has not been designed to avoid this problem.

The metals content of the coal ash also may be critical because metals can form deposits on waste heat boiler tubes of water wall tubes of the gasifiers and decrease tube life.

Manufacturing Processes

Today hydrogen is manufactured by four principal processes: (1) steam reforming of natural gas; (2) partial combustion of natural gas or oil with pure oxygen; (3) gasification of coal or coke with air or oxygen and steam; and (4) recovery of hydrogen from petroleum refinery gases or other cracking operations. Small amounts of hydrogen also are manufactured by electrolysis. These processes are discussed below.

Reforming. The reforming reaction is as follows:

$$CH_4 + H_2O \rightarrow 3H_2 + CO$$

The reaction is endothermic and requires the input of a large amount of heat. The heating gas requirement is about 80 percent of the process gas requirement. The reaction is carried out at temperatures near 800°C and at pressures ranging from atmospheric up to 500 psig.

Other reactions that proceed at the same time as the reforming reaction are:

$$CO + H_2O \rightleftarrows CO_2 + H_2$$

$$CH_4 \rightleftarrows C + 2H_2$$

$$2CO \rightleftarrows C + CO_2$$

The equilibrium composition depends upon the steam-to-gas ratio entering the reactor, the temperature, the pressure, and the quantity of inerts in the reaction mixture. One source[10] gives a general equation that is simplified to the following for the reforming of CH_4 with steam:

$$K = \frac{(N_{CO})(N_{H_2})^3}{(N_{CH_4})(N_{H_2O})}(Kv)\left(\frac{\pi}{N}\right)^2$$

where K is the equilibrium constant, N_{CO}, N_{H_2}, N_{CH_4}, and N_{H_2O} are the number of moles of each species present, Kv is the fugacity factor, π is the total pressure, and N is the total number of moles present. The equilibrium constant for the methane reforming reaction is as follows:

Temperature, °C	*Equilibrium Constant*
315	2.186×10^{-7}
425	2.659×10^{-4}
540	4.900×10^{-2}
650	2.679
760	0.6343×10^{2}
870	8.166×10^{2}
980	6.755×10^{3}

The final composition will also satisfy the carbon monoxide–shift reaction. The equilibrium constant for this reaction is:

Temperature, °C	*Equilibrium Constant*
315	31.4
425	9.03
540	3.75
650	1.97
760	1.20
870	0.819
980	0.604

To avoid carbon formation as indicated by the foregoing reactions, the steam-to-gas ratio must be kept high enough to favor the reforming and shift reactions in preference to those forming carbon.

A catalyst is required for this reaction. Requirements are: high activity, long life, good physical strength, and low cost. Such a catalyst can be prepared by mixing nickel oxide powder, alumina, and calcium aluminate cement, forming the mixture into the desired shapes (for example, $\frac{5}{8} \times \frac{5}{8} \times \frac{1}{4}$ inch Raschig rings) and then calcining the catalyst at a temperature of 650 to 1100°C.

Excellent properties are also obtained by impregnating a support material, such as alumina that has been fired at 1500°C with a nickel salt solution, and then firing the catalyst.[11]

Although silica and silica-bearing materials are very useful in making catalyst supports, they cannot be used for the shift catalyst because the silica volatilizes and migrates from the hotter zone to lower temperature zones downstream. Usually it deposits on the waste heat boiler tubes after the secondary reformer.

Normally, the nickel oxide is reduced to nickel and water by the hydrogen that is produced in the operation. According to U.S. Patent 3,827,987,[12] in some cases the reduced nickel can be reoxidized to nickel oxide when large amounts of steam and small amounts of H_2 are present:

$$Ni + H_2O \rightleftarrows NiO + H_2$$

U.S. Patent 3,827,987 claims that addition of 0.5 to 10 percent by volume of H_2 in the natural gas feed will keep the nickel in the reduced state, thus making it more active. Also, the hydrogen will retard the formation of nickel sulfide, thus avoiding poisoning of the catalyst:

$$NiS + H_2 \rightleftarrows Ni + H_2O$$

Poisons for the nickel catalyst are: sulfur, arsenic, chlorides or other halogens, phosphates. and copper or lead.[11] A 15 percent nickel catalyst is poisoned at 775°C, should the gas contain 0.005 percent sulfur. This is equivalent to reaction of all the nickel on the surface of the crystallites 1 micron in diameter. For lower operating temperatures, the amount required for poisoning is even lower. When using naphtha as a feedstock, 0.5 ppm of sulfur (w/w) in the naphtha is the maximum allowed concentration for operation at 775°C.

Carbon also can reduce the effectiveness of the catalyst. When conditions are favorable for the following reactions, carbon will

deposit:

$$CH_4 \rightarrow C + 2H_2$$

$$2CO \rightarrow C + CO_2$$

Thus, where there is insufficient steam, carbon will be deposited on the catalyst, reducing its effectiveness. Carbon will gasify with steam at 800°C or higher in the reformer:

$$C + H_2O \rightarrow CO + H_2$$

Operation and changes in operation should always be done to ensure sufficient steam over the catalyst; thus, if the feed rate is to be increased, the steam rate should be increased first. Conversely, if the feed rate is to be decreased, the natural gas rate should be decreased first.

A reforming reaction furnace usually consists of a number of 2- to 8-inch-diameter centrifugally cast chrome-nickel tubes (25% Cr, 20% Ni, 0.35–0.45% C) (Grade HK-40) about 10 to 40 feet long, mounted vertically in a refractory-lined furnace. The process gas and steam are fed downward over the catalyst and removed from the bottom of the tubes. The tubes are heated by burners located strategically in the sides, top, or bottom of the furnace. Close control of the heat is essential to ensure uniform temperature distribution throughout the tubes. To conserve natural gas for use as feed gas, some plants have converted from natural gas firing of the reformer to fuel oil firing.

During heating from atmospheric temperature to the reaction temperature (800°C), expansion of the catalyst tubes is considerable; thus, provision must be made for this growth. Usually, the tubes are suspended from springs, or counterweights are attached to the tops. Feed materials are introduced through flexible tubing.

At the high operating temperature in the reformer, severe tube damage can result from carryover of boiler feed-water salts.[13] The carryover of boiler feed-water salts can cause sulfidation and molten salt attack on the tubes, manifold, and transfer lines. Carbonization of methane and higher hydrocarbons will carburize most metals rapidly. Also, trace amounts of lead in the alloy will promote catastrophic oxidation, and burner fuels may contain sodium, vanadium, and sulfur compounds that will attack the outer surface to the tubes.

Some of the high-temperature transfer lines are protected by castable or firebrick linings, with the outer retaining piping water-cooled. The inner refractory liners also are protected by shrouds of Type 304 or 321 stainless steel.

Preheated steam and natural gas are mixed in a mole ratio of 3 to 8 prior to introducing the mixture into the catalyst tubes. Usually, heat exchange with the hot flue gases is sufficient to provide the preheat.

Steam to carbon ratios of 6 to 10 are claimed to be advantageous; such ratios result in less fouling of waste-heat boilers and give better reformer-catalyst and tube life.[14] Tube life is in the range of 6 to 12 years.

The hydrogen–to–carbon monoxide mole ratio for the product gas is usually 4 to 9; however, it can be increased if additional steam is used in the reaction. This will reduce the demand in the carbon monoxide–shift converters, which follow the secondary reformer. Steam sometimes is introduced after the reformer before the gas is fed to the carbon monoxide–shift converters. In the single-train ammonia plants, the natural gas is reformed in two steps. In the first step, the reaction takes place in the primary reformer in tubes suspended in a refractory-lined furnace. The large amount of endothermic heat is supplied by burning natural gas with air in the furnace. The heat flux in the tubes can be as high as 35,000 Btu/hr sq ft. The methane leakage is about 10 percent in the effluent dry gas, or about 60 to 65 percent of the feed methane is converted to synthesis gas.

For ammonia manufacture, it is necessary to introduce nitrogen into the system. This is done in the second step, as shown in Fig. 28.3, by mixing air with the primary reformer gas in a refractory-lined vessel that contains additional reforming catalyst. The mixing of the air and the primary reformer gas is critical, and specially designed burners are used. The temperature at the point of mixing will reach 1260°C because of the fast reaction of the

Fig. 28.3. Ammonia synthesis via natural gas reforming.

oxygen in the air; but then the endothermic reforming reaction takes place, and the gases cool to about 980°C. At this temperature, the methane concentration has been reduced to about 0.3 percent. The concentration is dependent upon the pressure, the temperature, and the quantity of nitrogen and steam present.

The catalyst used in this reforming step does not need to be as active as that in the primary reformer. Hence, the usual nickel concentration is about 15 percent, compared with 25 percent in the primary-reformer catalyst.

The hot gases from this section of the process are passed through waste heat boilers to generate the large amount of steam necessary to supply the process and drive the compressors downstream in the plant.

Partial Combustion. A flow diagram for ammonia production via partial combustion of natural gas is shown in Fig. 28.4. By burning natural gas with a limited quantity of oxygen, a synthesis gas containing approximately 2 moles of hydrogen for each mole of carbon monoxide can be produced:

$$CH_4 + \tfrac{1}{2}O_2 \rightarrow CO + 2H_2$$

The reaction is exothermic, reaching 1400°C, and must be carried out in a refractory-lined or water-cooled vessel. High temperatures favor the formation of carbon monoxide and hydrogen.[15] Pressure hinders the reaction, but pressures up to 1500 psi can be used and still produce a gas of low methane content. If steam is introduced into the reaction, the

Fig. 28.4. Ammonia synthesis via partial combustion of natural gas.

hydrogen–to–carbon monoxide ratio can be increased to over 2.

Preheating the feeds will reduce the requirements for oxygen and natural gas, as some of the natural gas must be burned to carbon dioxide in order to raise the gases to reaction temperature. The partial combustion reaction itself is not sufficiently exothermic to provide this heat.

The partial combustion of fuel oil or crude oil is practiced in many areas where natural gas is not available. In using crude oil, care must be taken to provide for the removal of sulfur compounds and ash-containing materials in the crude oil. These substances have been known to damage the refractory and, of course, to introduce sulfur into the product gas. In the case of crude-oil or fuel-oil partial combustion, the carbon content is much greater than in the case of natural gas, and special design considerations are necessary to produce a satisfactory gas.

Gas compositions from the partial combustion of various fuels are given in Table 28.7.

The nitrogen needed to produce the desired H_2/N_2 ratio usually is introduced later in the processing sequence.

Coal and Coke Gasification. An important process for the gasification of coal and coke is the Lurgi process, used in South Africa at the "coal to oil" plant and also used in Europe. This process requires the use of coke or a noncaking coal and oxygen; the product gas usually is high in methane. The reaction is carried out with a fixed bed of burning coal or coke, into which steam and oxygen are introduced. The oxygen requirement is 0.11 SCF O_2 per scf of synthesis gas when lignite is used, and 0.21 SCF O_2 per SCF of synthesis gas when anthracite is used.[16] This and other gasification processes are described in Chapter 17.

There are several processes for purifying synthesis gas made from coal, coke, or oil. The impurities usually include H_2S, CS_2, COS, ungasified carbon, and residual ash. Other factors that influence the selection of a purification process are the hydrogen–to–carbon monoxide ratio and the carbon dioxide content.

If the gas contains large amounts of sulfur, it is quite often the practice to remove this sulfur before the carbon monoxide–shift operation. There are a number of sulfur removal processes, namely, amine scrubbing, hot potassium carbonate scrubbing, and the Rectisol process. Ash removal and carbon removal usually are carried out before the gas is treated for sulfur or carbon dioxide removal and before the shift operation.

If the gas does not contain large quantities of sulfur and ash, the carbon monoxide–shift reaction is carried out directly after the synthesis operation. In some instances, it also is desirable to remove the carbon dioxide before the shift operation, as it has an adverse effect on the carbon monoxide–shift equilibrium. Carbon dioxide may be removed by scrubbing with water, monoethanolamine

TABLE 28.7 Product Gas Compositions from Various Feeds by Partial Combustion (percent by volume dry)

	Feedstocks					
Gas Components	*Nat. Gas*	*Light Naphtha*	*Heavy Fuel Oil*	*Vacuum Resid.*	*Propane Asphalt*	*Coal*
Hydrogen	61.80	52.09	46.74	44.80	43.69	34.39
Carbon monoxide	33.75	42.59	48.14	49.52	50.09	44.22
Carbon dioxide	3.10	4.88	3.80	4.16	4.45	18.64
Methane	1.00	0.30	0.30	0.30	0.30	0.38
$N_2 + A$	0.27	0.13	0.23	0.18	0.30	0.68
H_2S	—	0.01	0.76	1.00	1.12	1.46
COS	—	—	0.03	0.04	0.05	0.10
NH_3	—	—	—	—	0.13	
H_2/CO mol/mol	1.83	1.22	0.97	0.90	0.87	0.78

Fig. 28.5. Ammonia synthesis from refinery H_2.

(MEA), or hot potassium carbonate. The preferred process depends to some extent on the local economics and the particular conditions under which the synthesis gas is produced.

Hydrogen from Petroleum Gases. Petroleum gases from platforming operations, hydroformers, and butadiene plants may be processed to recover H_2. The gas from the platformer is particularly rich in H_2, containing as much as 90 to 95 percent. This gas usually is purified by low temperature fractionation or washing with liquid nitrogen (Fig. 28.5).

Electrolysis. In some areas hydrogen is obtained by electrolysis, but normally this is not economical. Electric power must be available at a very low cost such as that from hydroelectric power. In some cases, by-product hydrogen from the electrolysis of brine has been used to manufacture ammonia.

CARBON MONOXIDE SHIFT

Conversion of carbon monoxide to carbon dioxide and hydrogen with steam is necessary to make economical use of the raw synthesis gases produced by the foregoing processes. In this operation, the following reaction is promoted:

$$CO + H_2O \rightleftharpoons CO_2 + H_2$$

The reaction is exothermic, and high temperatures are unfavorable for complete conversion. The equilibrium is not affected by pressure.

The equilibrium constant for the reaction is as follows:

Temperature, °C	*Equilibrium Constant*
205	207
230	119
260	72.8
290	46.7
315	31.4
345	22.0
370	15.9

In recent years the shift operation has been carried out in two steps, one at a high temperature of 350 to 450°C, and one at a lower temperature of 200 to 300°C. The high-temperature shift reaction is carried out over an iron–chromium catalyst, which is not particularly sensitive to sulfur although high concentrations have been found to be detrimental. At the usual effluent temperature, the dry gas will contain about 3 percent carbon monoxide. To operate the low-temperature shift reaction, it is necessary to remove heat between the two catalyst beds.

High steam-to-gas ratios favor conversion of the carbon monoxide to hydrogen and a low temperature rise, resulting in a more favorable equilibrium. The normal carbon monoxide leakage is about 0.3 to 0.5 percent at an exit temperature of 240°C.

The low-temperature shift catalyst can be purchased in a reduced state. If it is not, it must be reduced before it is placed in operation. The cupric oxide must be reduced to metallic copper.

Experience in operation of the large single-train ammonia plants has shown that performance of the low-temperature shift catalyst is critical to good operating continuity. This catalyst contains copper and zinc and is very sensitive to sulfur and halogens.[17] Therefore, a guard catalyst frequently is used to protect it. The use of a small guard vessel ahead of the main low-temperature shift catalyst bed is valuable. The following operating practices have been found to be helpful:[18]

1. Dry reduction of the catalyst using natural gas or nitrogen as the carrier instead of steam.
2. The use of a low-temperature shift catalyst instead of zinc oxide as a guard catalyst.
3. Elimination of known sources of poisons.
4. A separate guard vessel.
5. Regular replacement of the guard catalyst.
6. Dry inlet temperature control rather than quenching.
7. Avoidance of an approach to the dew point of less than 22°C.

CARBON DIOXIDE REMOVAL

Following the shift conversion of CO with steam to CO_2 and H_2, the bulk of the CO_2 must be removed from the hydrogen and nitrogen. Carbon dioxide can be removed by scrubbing the gas under pressure with various solvents. Processes for CO_2 removal include:[19]

- Water
- Hot potassium carbonate—Benfield or Catacarb
- Monoethanolamine
- Sulfinol—sulfolane ($C_4H_8SO_2$) (diisopanol amine)
- Fluor solvent—propylene carbonate
- Rectisol—refrigerated methanol
- Vetrocoke (arsenical)
- Triethanolamine
- Selexol—dimethyl ether of polyethylene glycol
- Alkazid—potassium-*N*-methyl-aminopropionate
- Purisol—*N*-methyl-pyrrolidone (NMP)

Water Scrubbing

This system, once used widely, has not remained competitive. A large quantity of water is required to dissolve the carbon dioxide and special design is required to avoid excessive loss of nitrogen and hydrogen. Power recovery is important, and is achieved by use of properly designed turbines.

Hot Potassium Carbonate

In this system,[20] an aqueous solution containing 40 percent K_2CO_3 is recirculated at 118°C in the absorber. The CO_2 picked up is released in the stripper at atmospheric pressure by heating the solution to 230°C. An inhibitor generally is needed to reduce corrosion. Stainless steel equipment is used in many places. This system is characterized by its low steam requirements and its compatibility with the catalysts.

The system may be modified by adding materials that enhance the solubility of CO_2 and hence result in more complete recovery. Alkanolamines and As_2O_3 are typical additives.

Monoethanolamine (MEA)

A 15 to 30 percent solution of MEA in water is used to absorb the CO_2 under pressure. The solution then is regenerated by heating in a stripper in which the carbon dioxide is released. This system is characterized by good absorption properties at low pressure.

A corrosion inhibitor is used with this process. With the Amine Guard inhibitor, the MEA concentration in the circulating solution can be increased to 30 percent from a normal 20 percent. Hence, the circulating rate can be

decreased by 33 percent, and heat requirements are decreased 43 percent. Stainless steel is used in critical areas, where hot carbon dioxide-laden solutions are present.[21] It is reported that antifoam agents (e.g., certain silicones) are used in some cases with this process.

Sulfinol

The sulfinol process is a development by the Shell Development Corp., in Emoryville, California (U.S. Patent 3,352,631).[22] The solution consists of sulfolane ($C_4H_8SO_2$), which is a ring compound, diisopropanolamine, and water. This process is characterized by its high CO_2 retention under pressure and by low steam requirements for regeneration. A sidestream regenerator is required to remove by-products that build up in the system. See U.S. Patent 3,347,621.[23]

Mild steel is suitable in a number of places; however, stainless is preferred where CO_2 concentrations are high.

The process must be carefully engineered to protect the downstream nickel catalysts from the sulfur, which may carry over from the absorber. Likewise, the CO_2 that is produced must be protected from entrainment of the solvent and contamination with sulfur.

Propylene Carbonate

This process was developed primarily for the removal of high concentrations of CO_2 from natural gas. More than 11 plants apply the Fluor solvent, propylene carbonate, process. Regeneration of the solvent is accomplished by staged pressure reductions with flashing at each stage, rather than by heating. This process features power recovery by using some of the flashed gases to drive turbines.

The absorber operates at a relatively low temperature, made possible by the refrigerating effect of the released CO_2; and some auxiliary refrigeration is needed. Mild steel is suitable for construction in most areas of the system. If the CO_2 is needed for urea manufacture, it can be released at 100 to 200 psi.[24,25]

Rectisol—Refrigerated Methanol

This system is licensed by Lurgi GmbH, and more than 65 units are in operation. The methanol is cooled to $-50°C$, at which temperature the vapor pressure is less than 1 mm Hg. The Rectisol process, described in U.S. Patent 2,863,527,[26] is used chiefly for the following applications:

1. "Complete and simultaneous removal of CO_2, H_2S and organic sulfur, NH_3, hydrogen cyanide, gumformers, higher hydrocarbons and other impurities from coal gasification crude gas to produce gas for synthesis or pipeline transmission."
2. "Removal of H_2S, COS, and CO_2 from reformed gas, in particular from gas produced by partial oxidation of hydrocarbons, to yield synthesis gas."
3. "Integration of gas purification with low temperature plants (liquefaction and fractionation) for the complete removal of moderate contents of acidic components."

The system can be designed to produce a stream rich in H_2S for processing in a Claus unit to recover sulfur and meet emission standards.

Giammarco—Vetrocoke

This process is based upon the absorptive power of potassium carbonate solution activated by glycine (arsenic in older plants). Both H_2S and CO_2 are removed. By varying process conditions, the CO_2 level can be reduced as low as necessary, and CO_2 with a purity of 99 percent can be produced. The solution is reported to be noncorrosive to steel. More than 200 plants are in operation, and newer plants have substantially reduced energy requirements.[27]

FINAL PURIFICATION

Before the ammonia synthesis gas ($3H_2/1N_2$) is admitted to the ammonia synthesis step, essentially all of the carbon monoxide and

carbon dioxide must be removed, as they are temporary poisons for the synthesis catalyst. Also, the presence of carbon dioxide in the makeup gas causes difficulty in the synthesis loop because it forms carbamates and ammonium carbonate, which are solids that can damage compressors and valves.

Methanation

This is the most popular process for final purification or removal of the last traces of carbon monoxide and carbon dioxide. This process is the reverse of the methane reforming process:

$$CO + 3H_2 \rightleftarrows CH_4 + H_2O$$

$$CO_2 + 4H_2 \rightleftarrows CH_4 + 2H_2O$$

The large excess of hydrogen present and a lower temperature favor the complete removal of carbon oxides, usually to less than 10 ppm.

In this process, the synthesis gas, which has been scrubbed to remove the bulk of the CO_2, is preheated to about 315°C and then passed over a nickel catalyst (NiO and Al_2O_3). Ruthenium also is effective as a catalyst. The space velocity depends upon the pressure of the process. At 450 psig the space velocity (SCF of gas per hour per cubic foot of catalyst) is about 5000 hr^{-1}. At lower pressures a space velocity of 2000 hr^{-1} is achieved.

The catalyst life is normally one to three years. Sulfur and halogens are poisons and shorten the life.

Equilibrium constants[28] for the two reactions are:

$T°C$	$\frac{(P_{CH_4})(P_{H_2O})}{(P_{CO})(P_{H_2})^3}$	$\frac{(P_{CH_4})(P_{H_2O})^2}{(P_{CO_2})(P_{H_2})^4}$
200	0.21547×10^{12}	0.94748×10^{9}
260	0.45626×10^{9}	0.62706×10^{7}
320	0.32635×10^{7}	0.11001×10^{6}
380	0.56011×10^{5}	0.38882×10^{4}
440	0.18550×10^{4}	0.23282×10^{3}
500	0.10219×10^{3}	0.20997×10^{2}

Both reactions are exothermic; the heats of reaction are ($\Delta H_{25°C}$) −49.27 and −39.44 kcal/mole of carbon monoxide and carbon dioxide, respectively.

Nitrogen Wash Operation

The last traces of carbon monoxide are removed by liquid nitrogen scrubbing in some plants, generally those using a partial oxidation process for generation of the synthesis gas. The nitrogen is obtained from the air-separation plant used for oxygen production. In the nitrogen-wash operation, the gas must be thoroughly dried and the carbon dioxide removed before the gas is cooled to the point where it can be scrubbed with liquid nitrogen. Otherwise, the moisture and the carbon dioxide would freeze out in the equipment, causing a number of problems. Hence, the gas must be passed through silica gel driers and a caustic wash to remove the water and carbon dioxide. At this point, the gas can safely be cooled to the liquid-nitrogen temperature and scrubbed to remove the carbon monoxide, methane, and argon.

The overhead gases from the wash tower usually contain about 90 percent hydrogen and 10 percent nitrogen, thus requiring the addition of more nitrogen to give the required ratio for ammonia synthesis. The gas so produced is suitable for ammonia synthesis.

Copper Ammonium Carbonate Scrubbing

In some instances, it is preferable to use copper scrubbing to remove the last traces of carbon monoxide in the ammonia synthesis gas. This process usually is carried out at pressures in the range of 1000 to 1500 psi, which are intermediate pressures of the high pressure ammonia synthesis feed compressors. The copper liquor is circulated countercurrent to the synthesis gas, picking up the carbon monoxide as a copper ammonium carbonate–carbon monoxide complex and producing a relatively carbon monoxide–free gas. Methane and argon, however, are not removed; they pass on with the hydrogen to the ammonia synthesis units.

The copper liquor is regenerated by reducing the pressure on the liquid, which allows the carbon monoxide complex to decompose. The carbon monoxide is separated from the copper liquor, and the liquid is recirculated to the scrubbing operation.

Selective Oxidation of Carbon Monoxide

This process is based on the fact that over a platinum catalyst at 120 to 160°C carbon monoxide will be selectively oxidized to carbon dioxide.[29] Excess oxygen is converted to water by hydrogen.

$$CO + \tfrac{1}{2}O_2 \rightarrow CO_2$$

$$H_2 + \tfrac{1}{2}O_2 \rightarrow H_2O$$

The disadvantage of the process is that a second scrubbing step is necessary to remove the CO_2 formed by oxidation of the carbon monoxide.

Cryogenic Purifier

In the C. F. Braun process for ammonia manufacture, excess air is added in the second reformer.[30] The excess nitrogen thus introduced must be removed prior to the synthesis step to avoid excessive loss of hydrogen and excessive compression costs. At the same time that the excess nitrogen is condensed and removed, the remaining traces of carbon monoxide, the methane, and most of the argon are removed, leaving a gas comparable to that produced by the nitrogen wash operation.

U.S. Patent 3,584,998[31] claims the primary reformer can be eliminated by adding more excess air, preheating the air, natural gas, and steam, and carrying out the entire reforming step in a single-bed vessel.[32] In this process modification, the feed to the purifier would contain 60 to 70 percent hydrogen, 30 to 40 percent nitrogen, 2 to 3 percent methane, and about 0.5 percent argon.

COMPRESSION

With the advent of large-capacity plants (over 600 tons/day), it has been possible to design centrifugal compressors for the services required. The pressure level has also dropped to about 2000 psi to favor compressors successfully used by industry. However, more and more machines are now being built for higher pressures, up to 5000 psig. One of the advantages of the centrifugal compressors is that the need for lubricating oil is minimal; oil contamination causes difficulty in the synthesis loop.

The centrifugal compressors are driven by steam turbines using high pressure steam generated in the plant waste heat boilers heated by process gas leaving the secondary reformer and the flue gases from the primary reformer. Design of the compressors is critical from the standpoint of efficiency and economy. About 25 percent of the downtime of large plants results from compressor troubles.

About 20 percent of the energy used in ammonia manufacture goes into compression (synthesis gas, air, and refrigeration). A study of new alternatives suggests that gas turbines and more efficient steam systems (higher temperatures and reheat) should be considered.[33]

For smaller plants, under 600 tons/day, reciprocating compressors still are used. They often contain several services on a common drive shaft: air compression, natural gas compression, synthesis gas compression and recycle, and ammonia compression for regrigeration.

AMMONIA SYNTHESIS

Ammonia synthesis is carried out at pressures ranging from 2000 to 10,000 psi. The preferred pressure depends largely on the quality of the synthesis gas and certain other conditions, such as production requirements per converter.

Quartulli et al.[34] published a detailed study of the factors to be considered in selecting the proper synthesis pressure. The reaction equilibrium, reaction rate, and condensation of ammonia all are favored by high pressure. Mechanical problems and cost considerations also are considered. In general, the study concluded that operation at 2100 psi is more economical than at 3300, 4700, or 6100 psi. The following were some of the factors considered: catalyst volume, recycle flow, horsepower for makeup-gas compression, recycle-gas flow, refrigeration, equipment cost, and equipment reliability.

The ammonia synthesis reaction is:

$$N_2 + 3H_2 \rightarrow 2NH_3 \text{ exothermic}$$

Lower temperatures favor higher equilibrium concentrations of ammonia; the rate of reaction at low temperature is so slow that the process is uneconomical at temperatures below about 400°C. Usually a conversion of about 15 to 30 percent per pass is obtained when the temperature is in the range of 400 to 600°C.

The heat of reaction for the synthesis of ammonia is given by the following equation:

$$\begin{aligned}\Delta H(\text{cal/mole}) = -\left[0.54526 + \frac{840.609}{T} + \frac{459.734 \times 10^6}{T^3}\right]p \\ - 5.34685T \\ - (0.2525 \times 10^{-3})T^2 \\ + (1.69167 \times 10^{-6})T^3 \\ - 9157.09\end{aligned}$$

where T is °K and p is in atmospheres.[35]

Reaction Rate

The rate of reaction over the synthesis catalyst depends upon the type of catalyst, rate constants for synthesis and decomposition, and the partial pressures of nitrogen, hydrogen, and ammonia. Temkin and Pyzhev[35] propose the following equation:

$$W = K_1 P_{N_2}\left(\frac{(P^3_{H_2})}{(P^2_{NH_3})}\right)^{\alpha} - K_2\left(\frac{(P^2_{NH_3})}{(P^3_{H_2})}\right)^{\beta}$$

where W is the reaction rate; K_1 and K_2 are the rate constants for synthesis and decomposition, respectively; P_{H_2}, P_{N_2}, and P_{NH_3} are partial pressures; α, β are constants. The constant α has been found to be in the range of 0.5 to 0.75, and β in the range of 0.4 to 0.3; $\alpha + \beta = 1.0$. Tempkin and others established this equation on the basis that the rate-determining steps are the chemisorption of nitrogen in a diatomic form and the combustion of a chemisorbed nitrogen molecule with a hydrogen molecule.

Fig. 28.6. Ammonia production rate as a function of pressure.

At high pressure and with some ammonia in the gas, the effect is to have only one rate-determining step.[36] Figure 28.6 gives some indication of the production rate as a function of pressure. Most systems are designed for 2000 to 5000 psi.

The gas composition for ammonia synthesis also is critical, relative to the cost of ammonia. High concentrations of inert materials such

In design of the equipment, it is necessary to take into account the solubility of the various synthesis loop gases in liquid ammonia. The solubilities[37] of the gases at normal pressure and temperature are, in cubic centimeters of gas at normal pressure and temperature per gram of liquid ammonia per atmosphere partial pressure:

Gas	*5°C*	*30°C*
Hydrogen	0.095	0.112
Nitrogen	0.117	0.126
Argon	0.151	0.161
Methane	0.270	0.392

Note that all four gases are more soluble at the higher temperature.

The optimum concentration of inerts will depend on the temperature and pressure of the operation and the concentration of inerts

in the makeup gas. The solubilities of argon, methane, hydrogen, and nitrogen in liquid ammonia determine the amounts of these gases that are removed from the system with the liquid ammonia. The designer has some control over this purge. The conditions of synthesis can be selected to give optimum purge rates once solubility, makeup gas composition, and degree of conversion are determined.

The production rate of ammonia is dependent upon the recycle rate for the gases in the loop, the degree of conversion attained in the catalyst, and the efficiency with which the ammonia is removed. Table 28.8 gives some ranges for the saturated-vapor ammonia concentration in a $3H_2:1N_2$ gas mixture over liquid ammonia.

Ammonia liquid viscosity[38] may be calculated from the following equation:

$$\mu = bT^m$$

where:

μ = viscosity in centipoises, cP
$b = 6.301 \times 10^8$ correlation constant
$m = -3.921$
T = temperature, °K

The value of μ at −50°C is 0.317 cP.

TABLE 28.8 Ammonia Concentration Over Liquid Ammonia

Pressure (atm)	*Temperature (°C)*	*NH_3 in Gas %*
50	−20	5.7
	0	10.0
	20	19.0
	30	25.5
100	−20	3.25
	0	5.9
	20	11.0
	30	15
200	−20	2.05
	0	3.95
	20	7.5
	30	10.04
300	−20	1.62
	0	3.25
	20	6.4
	30	8.8
700	−20	1.19
	0	2.35
	20	4.90
	30	6.80

Converter Details

Two relatively new converter arrangements are shown in Figs. 28.7 and 28.8. They both contain an internal heat exchanger, and the bypass inlet is for temperature control purposes. As in most converters, the cold inlet gas passes between the catalyst basket and the shell to prevent overheating of the shell and hydrogen embrittlement.

Figure 28.7 shows a radial-flow converter, a design by Topsoe. The main feature of this converter is the provision of a much greater

Fig. 28.7. Topsoe radial-flow converter. (*"Nitrogen," Sept. 1964, The British Sulfur Corporation.*)

Fig. 28.8. Horizontal ammonia converter, Kellogg design. (Hydrocarbon Processing, *pp. 115–122, Dec. 1978, copyright 1978 by the Gulf Publishing Company.*)

Fig. 28.9. Circulating system for ammonia synthesis.

cross section to the flow of gas, thereby reducing pressure drop. This low pressure drop makes possible the use of smaller catalyst particles with some increase in activity per unit volume.

Kellogg has introduced a horizontal converterter, as shown in Fig. 28.8. This arrangement gives a large cross-sectional area for the flow of gas while still using a shell that permits full closure. A small catalyst size is thus permitted, with an accompanying reduction in volume.

In addition to the care that must be exercised with respect to materials of construction, it is necessary to design for heat removal. This usually is accomplished by internal heat exchangers in the converter. The incoming gas is heated by the exit gas, which is thereby cooled. A temperature rise of about 150°C is used in the design of a converter.

A circulating system for ammonia synthesis is shown in Fig. 28.9. Control of the gas composition is important in the synthesis loop. The means of achieving this is revealed in U.S. Patent 2,894,821.[39] The hydrogen-to-nitrogen ratio must be kept close to 3, and the concentration of inert materials, methane and argon, must be limited to levels that will

permit a satisfactory partial pressure of the H_2 and N_2. In the process described in the patent, the circulating gas in the synthesis loop is analyzed for hydrogen and nitrogen, and the inlet air flow to the secondary reformer is adjusted to maintain the ratio of hydrogen to nitrogen at 3.0.

Control of many modern ammonia plants is accomplished via gas analyzers and on-line computers to adjust conditions for optimum performance. The control strategy includes adjustments to the H_2/N_2 ratio, purge rate, and reformer furnace control. Heat recovery during regeneration of the catalysts also is included.

Catalysts for ammonia synthesis are made by mixing natural or synthetic magnetite and promoters, and melting the mixture in an electric furnace. The cooled mass is broken up, sized, and in some cases pre-reduced before charging to the converter. The particle size is 2 to 10 mm, and the bulk density is in the range of 2.5 to 2.9 kg/liter. A typical chemical composition is as follows: 68.4 percent total Fe, 3.16 percent Al_2O_3, 0.56 percent MgO, 0.50 percent SiO_2, 3.54 percent CaO, and 0.58 percent K_2O.

The catalyst must be reduced to the metallic constituents before it is placed in operation.[11] For a multi-bed quench converter, the synthesis gas is circulated in the loop at operating pressure, heated at 50°C per hour to 340°C and then at 10°C per hour until the desired temperature is reached. The water content of the exit gas should be about 5000 to 10,000 ppm. As ammonia is made, it will reduce the heat input that was required at the start. When temperatures rise of their own accord, the circulation rate can be increased. Temperatures should be held below 500°C.

A highly active catalyst has been described in U.S. Patent 3,992,328.[40] The reduced iron catalyst is treated with an aqueous solution of cerium nitrate and re-reduced after packing in the converter. Exit ammonia concentrations from a laboratory test unit are about 3 to 4 percentage points higher in the exit gas for this Lummus catalyst.

Innovations for the ammonia synthesis loop include the recovery of argon and hydrogen, with recycle of the hydrogen, from the synthesis-loop purge gas. With a system designed for argon recovery, streams rich in methane and hydrogen can be obtained.[41] The purge gas is cooled and is washed with liquid nitrogen, and then the liquid nitrogen stream is fractionated in a two-column train to give liquid argon and methane. A stream of H_2 and N_2 is recycled to the loop.

Environmental control is essential to meet EPA demands. It has been estimated that for a plant consuming 37 million Btu/ton of ammonia produced, environmental control will require 0.7 million Btu, and environmental costs will be 1.1 percent of the operating costs. Discharge streams that require processing to eliminate undesirable contaminants include: (1) the regeneration acid and base for boiler feed-water treatment units, (2) the regeneration steam for natural gas desulfurization catalysts, (3) the low temperature shift condensate (ammonia, amines, and methanol), and (4) CO_2 removal process regeneration effluents.

The ammonia process can be combined with a urea process, and some operating and investment efficiencies can be realized.[42] The synthesis gas generated by conventional means is mixed with liquid ammonia before the carbon dioxide is removed. This mixture of ammonia and carbon dioxide is synthesized to urea, and the off-gas from this operation is used for ammonia synthesis. The carbon dioxide–removal step thus is eliminated, and the compression of carbon dioxide for urea is avoided.

Modern Single-Train Ammonia Plants

A large number of licensors are available, and competition has forced licensing companies to optimize and economize the plants. Most ammonia is produced from plants using natural gas as the feed raw material. Energy conservation and environmental control are major concerns.

Modern plants licensed by Bechtel and Kellogg exemplify the move toward large, computer-controlled systems. Kellogg technology has been used in plants responsible

Fig. 28.10. Sheritt Gordon plant at Fort Saskatchewan, Alberta is the first commercial application of Kellogg's reduced energy design and the most energy-efficient ammonia plant in operation today.

Fig. 28.11. Control room, Kellogg ammonia plant.

for about one-half the world's new fertilizer ammonia output.[43]

An example of the most modern and most efficient plant is the 1100-ton/day Kellogg plant in Canada,[44] shown in Fig. 28.10. This reduced energy process employs advanced distributive control, shown in Fig. 28.11. This plant was able to reduce the energy consumption to less than 25 million Btu/short ton of ammonia from the 32 million Btu of conventional plants.

Bechtel also licensed an efficient ammonia

plant of 1600 metric tons/day in 1991 with net energy consumption of 28.6 million Btu/metric ton of ammonia.[45]

Typical net energy consumption is:

Natural gas feed	22.37
Natural gas fuel	8.63
Natural gas total	31.00
Electric power	0.81
Gross energy consumption	31.81
Export steam credit	−3.22
Net energy consumption	28.59

USES OF AMMONIA

The uses for ammonia are given in Tables 28.9 and 28.10. By far the greatest consumption is in the manufacture of fertilizers or as a direct application for fertilizer.

TABLE 28.9 Uses of Ammonia[a]

	1,000 Short Tons
Ammonium nitrate	1,600
Direct fertilizer application	4,880
Urea	4,230
Ammonium sulfate	600
Nitric acid, industrial	2,300
Ammonium phosphates	3,150
Acrylonitrile	600
Others	1,350
Total	18,710

TABLE 28.10 Industrial Uses of Ammonia

Industry	*Usage*
Explosives	Nitrates, dynamite, azides
Plastics	Nitrocellulose, urea-formaldehyde, melamine, methylmethacrylates, paints
Metallurgy	Bright annealing of steel, dry reducing gas
Pulp and paper	Ammonium bisulfate, melamine
Rubber	Aniline, acrylonitrile, polyurethanes, chemical blowing agents (for foam rubber)
Textiles	Nylons, acrylonitrile, terephthalates
Foods	Amino acids, sodium nitrite, and nitric oxides
Drugs	Vitamins, nitrofurans, nitrous oxide
Miscellaneous	Refrigerant, detergents, insectides, nitroparaffins, hydrazine, air pollution control

STORAGE AND TRANSPORT

Ammonia usually is stored in large 20,000-ton atmospheric-pressure tanks at a temperature of −28°F. With this system, the ammonia vaporized by heat in-leakage must be compressed and condensed, and then returned to the system.

Some of the low-pressure tanks are built with a double wall; this adds some protection against leakage and also provides a means for retaining insulation.

Normally, ammonia is transported by barge, ship, or tank cars to terminals located at strategic points in the consuming agricultural areas. Some ammonia is transported by interstate pipelines.

Ammonia Price and Process Licensers

List price for ammonia, f.o.b. Gulf Coast in tank cars was reported at $110/short ton in April 1991.[46] The price follows energy prices very closely and can vary by as much as $30/ton over short periods of time.

Licensers in 1990 for ammonia plant technology were:[47]

Ammonia Casale	M. W. Kellogg
Bechtel	KTI
C. F. Braun	Linde
CECP	Lurgi/Shell
ChemieLinz/Chemserv	Nippon Kokan
Fish	Polimex-Cekop
Haldor Topsoe	Pritchard
Howe-Baker	Quimigal
ICI Chemicals	Uhde

NITRIC ACID

Nitric acid has been known since the thirteenth century. Glauber devised its synthesis from strong sulfuric acid and sodium nitrate; however, it was Lavoisier who showed that

nitric acid contained oxygen, and Cavendish showed that it could be made from moist air by an electric spark.

In the oldest methods used, Chile saltpeter ($NaNO_3$) was reacted with concentrated sulfuric acid in heated cast iron retorts; the evolved nitric acid vapors were condensed and collected in stoneware vessels.

Today, nitric acid is made by the Ostwald process. In this process, ammonia is oxidized with air over a precious metal catalyst at 800 to 950°C and at pressures from atmospheric up to 120 psig. The overall reaction is:

$$NH_3 + 2O_2 \rightarrow HNO_3 + H_2O - 98.7 \text{ kcal (evolved)}$$

The concentration of nitric acid usually is about 60 percent with conventional equipment. If higher concentrations are desired, special equipment or processes are required.

Nitric acid also can be produced by the high temperature combination of nitrogen with oxygen in the air. A method that received considerable attention in the 1950s is the so-called Wisconsin process, a thermal process for combining atmospheric oxygen and nitrogen. Although a 40-ton/day plant was built, it could not compete economically with the conventional ammonia-oxidation route.

Nitric acid is generally a light-amber liquid because of dissolved NO_2; however, the pure acid is colorless, strongly hygroscopic, and corrosive. It is a strong oxidizing acid. It boils at 78.2°C, freezes at −47°C, and forms a constant-boiling mixture with water (68% nitric acid by weight), which boils at approximately 122°C at atmospheric pressure. White fuming nitric acid usually contains 90 to 99 percent by weight HNO_3, from 0 to 2 percent by weight dissolved NO_2, and up to 10 percent by weight water. Red fuming acid usually contains about 70 to 90 percent by weight HNO_3, 2 to 25 percent by weight NO_2, and up to 10 percent by weight water. Tables 28.11 and 28.12 give properties of nitric acid.

Chemistry of Ammonia Oxidation

The oxidation of ammonia catalytically is an extremely rapid heterogeneous reaction. The platinum, platinum–rhodium, or platinum–rhodium–palladium catalyst in the form of a fine mesh gauze is very selective and very active under optimum conditions. Side reactions can occur resulting in a loss of product.

TABLE 28.11 Physical Properties of Nitric Acid Solutions[a]

wt. % HNO_3	Density (at 20°C), g/cm³	Freezing point, °C	Boiling point, °C	Specific heat (at 20°C) J/(g·K)[b]	Partial press. (at 20°C) Pa[c] HNO_3	Partial press. (at 20°C) Pa[c] H_2O	Viscosity (@20°C) mPa·s (=cP)	Thermal conduct. (@ 20°C) W/(m·K)[d]
0.0	0.99823	0	100.0	4.19		2333	1.0	0.61
10.0	1.0543	−7	101.2	3.73		2266	1.1	0.57
20.0	1.1150	−17	103.4	3.39		2026	1.2	0.54
30.0	1.1800	−36	107.0	3.18		1760	1.4	0.50
40.0	1.2463	−30	112.0	3.01		1440	1.6	0.47
50.0	1.3100	−20	116.4	2.85	27	1053	1.9	0.43
60.0	1.3667	−22	120.4	2.64	120	653	2.0	0.40
70.0	1.4134	−41	121.6	2.43	387	347	2.0	0.36
80.0	1.4521	−39	116.6	2.22	1400	120	1.9	0.35
90.0	1.4826	−60	102.0	1.97	3600	27	1.4	0.31
100.0	1.5129	−42	86.0	1.76	6000	0	0.9	0.28

[a]Sources: *International Critical Tables*, 1928; Perry, R. H., and Chilton, C. H., *Chem. Engrs. Handbook*, 1973.[125–127]
[b]To convert J/(g·K) to cal/(g·°C), divide by 4.184.
[c]To convert Pa to atm, divide by 1.013×10^5.
[d]To convert W/(m·K) to Btu/(h·ft·°F), divide by 1.7307.

TABLE 28.12 Thermodynamic Properties of Nitric Acid and Its Hydrates[a]

	HNO_3	$HNO_3 \cdot H_2O$	$HNO_3 \cdot 3H_2O$
Nitric acid, wt. %	100.0	77.77	53.83
Freezing point, °C	−41.59	−37.62	−18.47
Heat of formation (at 25°C), kJ/mol[b]	−173.35	−472.07	−888.45
Free energy of formation (at 25°C), kJ/mol[b]	−79.97	−328.29	−810.99
Entropy (at 25°C), kJ/(mol·K)[b]	155.71	217.00	347.17
Heat of fusion, kJ/mol[b]	10.48	17.52	29.12
Heat of vaporization (at 20°C), kJ/mol[b]	39.48		

[a]Source: Rossini, F. D., and Coworkers, Nat. Bur. Stand. Circ. 500, 53 (1952).
[b]To convert J to cal, divide by 4.184.

When ammonia is oxidized in air over a catalyst, the following reactions take place:

$$NH_3 + \tfrac{3}{4}O_2 \rightarrow \tfrac{1}{2}N_2 + \tfrac{3}{2}H_2O \quad (1)$$

$$NH_3 + O_2 \rightarrow \tfrac{1}{2}N_2O + \tfrac{3}{2}H_2O \quad (2)$$

$$NH_3 + \tfrac{5}{4}O_2 \rightarrow NO + \tfrac{3}{2}H_2O \quad (3)$$

$$NH_3 + \tfrac{7}{4}O_2 \rightarrow NO_2 + \tfrac{3}{2}H_2O \quad (4)$$

$$NH_3 \rightarrow \tfrac{1}{2}N_2 + \tfrac{3}{2}H_2 \quad (5)$$

The equilibrium constants for these reactions are given in Table 28.13.

Equation (3), for the formation of NO can be promoted by the selection of the right conditions and the proper catalyst. The nitric oxide yield increases as the temperature is increased. At atmospheric pressure, with 80 mesh gauze made from 0.003 inch wire (8% rhodium–92% platinum), 900°C, and 100 lb of NH_3 per troy oz. of catalyst per day, the conversion efficiency is about 99 percent.[48]

As converter pressure increases, efficiencies for conversion of NH_3 to NO fall off.[49] At one atmosphere gauge, conversion decreases to 97 to 98 percent; at 4 atmospheres there is 95 to 96 percent conversion, and at 10 atmospheres 92 to 93 percent. Increased pressure does, however, increase the amount of ammonia that can be oxidized per unit of catalyst and increases the absorption efficiency of NO_2 in water. In modern plants these competitive pressure effects have been optimized by operating the converter at about 72 psig and compressing the gases for absorption to 160 psig. The space velocity should be kept high in order to avoid decomposition of ammonia in the feed gas on the reactor walls by reaction (1) above. Keeping the temperature of the reactor walls low minimizes this reaction.

Also, contact of NO_2 with NH_3 must be minimized to avoid their reaction ($3NO_2 + 4NH_3 \rightarrow \tfrac{7}{2}N_2 + 6H_2O$). This is done by keeping the catalyst bed shallow and the velocity high enough to avoid back mixing.

TABLE 28.13 Equilibrium Constants for Reactions of Ammonia Oxidation with Air*

Equilibrium Temp. °F	*Kp, atm, for Formation of: (Equation No.)* $N_2(1)$	$N_2O(2)$	$NO(3)$	$NO_2(4)$	$H_2(5)$
440	7.07×10^{34}	4.44×10^{28}	1.13×10^{26}	1.43×10^{28}	3.33
800	2.61×10^{25}	2.69×10^{20}	2.11×10^{19}	5.04×10^{19}	1.11×10^{2}
1160	1.49×10^{20}	7.36×10^{15}	3.8×10^{15}	9.94×10^{14}	8.84×10^{2}
1520	6.71×10^{16}	9.12×10^{12}	1.54×10^{13}	—	3.21×10^{3}
1860	3.18×10^{14}	8.85×10^{10}	3.36×10^{11}	—	8.14×10^{3}

*Source: Sauchelli, "Fertilizer Nitrogen," ACS Monograph Series 161.[48]

The effect of reactor residence time on nitric oxide yield[48] is indicated by the following:

Reactor Residence Time, Seconds	*Nitric Oxide Yield, Percent NH_3 Oxidized*
0.28	82.1
0.11	85.7
0.061	90.2
0.023	91.8

The concentration of ammonia in the feed gas is kept below 14 percent because of the lower spontaneous inflammability of the mixture.

The desired reactions taking place in the process are:

$$NH_3 + \tfrac{5}{4}O_2 \rightarrow NO + \tfrac{3}{2}H_2O + 69.9 \text{ kcal evolved/g mol}$$

$$\tfrac{3}{2}NO + \tfrac{3}{4}O_2 \rightarrow \tfrac{3}{2}NO_2 + 13.6 \text{ kcal evolved/g mol}$$

$$\tfrac{3}{2}NO_2 + \tfrac{1}{2}H_2O \rightarrow HNO_3 + \tfrac{1}{2}NO + 8.35 \text{ kcal evolved/g mol}$$

The oxidation of the ammonia to NO takes place on the catalyst gauze, whereas the oxidation of the NO to NO_2 takes place when the gases are cooled by process equipment before and in the absorber itself. Equilibrium constants for the reaction

$$NO + \tfrac{1}{2}O_2 \rightarrow NO_2$$

are as follows:

Temp. °F	*Equilibrium Constant,* K_p; $atm^{-0.5}$*
35	9.42×10^6
80	1.19×10^6
170	4.61×10^4
260	3.95×10^3
350	7.24×10^2
530	35.2

*See reference 48.

Note the negative temperature coefficient of the reaction. A low temperature favors nitrogen oxide oxidation reaction as well as the absorption.

Processes

The first processes for oxidation of ammonia operated at atmospheric pressure. Over the years both high pressure and dual pressure processes have been used. Early processes used are given in Table 28.14.

Today the dual pressure process is widely used. In this process, the oxidation of ammonia is carried out at 72.5 psig and the absorption of the oxidation products at 160 psig. The low pressure for oxidation promotes high efficiencies of conversion and avoids high losses of the precious metal catalyst. Figure 28.12 shows a flow chart for a dual pressure plant using the Grande Paroisse process. Amounts of ammonia, catalyst, and energy consumed for two processes are estimated as follows:

	Per Ton of 100% HNO_3	
Material Consumed	*Grand Paroisse**	*SABAR***
Ammonia	280 kg	281
Platinum	90 mg	60
Electric Power	9 Kwh	335
Cooling Water	100 m^3	180
Steam, excess	900 Kg	0

**Hydrocarbon Proc.* 1983.[50]
***Hydrocarbon Proc.* 1989.[51]

In the Grand Paroisse dual-pressure process, the ammonia is vaporized, preheated, and filtered before mixing with the air, which is filtered and compressed to approximately 4 atmospheres. Part of the air goes to the acid bleacher, in which the air is used to strip

TABLE 28.14 Early Processes for Nitric Acid Manufacture

Process	*Temperature °F*	*Pressure*	*Acid Strength %*
Low Pressure	800	Atmos.	50–52
Medium Pressure	850	40 psi	60
High Pressure (Du Pont)	950	120 psi	60
Pintsch. Bamag	850	Atmos.	98–99

Fig. 28.12. Dual pressure nitric acid process. (Hydrocarbon Processing, *p. 119, Nov. 1983, copyright 1983 by the Gulf Publishing Company. By permission.*)

dissolved nitrogen oxides out of the nitric acid. The mixture of air and ammonia passes down over the platinum–rhodium catalyst. The converter effluent is cooled in a tail-gas heat exchanger and other exchangers before entering the NO compressor where the pressure is raised to 8 to 13 atmospheres. Both the air compressor and the NO compressor are driven by a steam turbine.

The compressed converter gas is cooled in a series of exchangers and enters an absorption tower where water is added. The tail gas from the absorber is heated through three exchangers and then enters a power recovery–expansion turbine. The absorber tower nitric acid is fed to the top of the bleacher, flowing down countercurrent to the air. The air, which leaves at the top of the bleacher, combines with the converter effluent and is recycled to the absorption tower. The tail gas from the power recovery turbine contains less than 150 ppm nitrogen oxides.

In the first stage of a nitric acid process where ammonia and air are mixed, it is important to hold the proper ratio of constituents. Usually a mixture of about 10 percent ammonia and 90 percent air is employed. At this point, it also is important to avoid the decomposition of ammonia on the converter walls. Aluminum is superior to mild steel in this respect. At 350°C the rate of decomposition is 300 times faster on mild steel than on aluminum. Ammonia is decomposed 70 times faster on stainless steel than on aluminum.[52] However, by keeping the preheat temperature low, satisfactory yields can be obtained with mild steel. The temperature level used should be below 200°C.

In the reaction that takes place over the catalyst bed, the rate-limiting step is the diffusion of ammonia to the catalyst surfaces. With a high gas-rate through the gauze, some of the ammonia will not reach the catalyst surfaces; it reacts instead with the nitric oxide product to give elemental nitrogen:

$$4NH_3 + 6NO \rightarrow 5N_2 + 6H_2O$$

The linear flow of gas across the catalyst thus is important and will depend on the arrangement of the catalyst. Linear flow rates of 1 to 3 ft/sec have given high efficiency. These rates are for gas at standard temperature and pressure. Without selection of proper flow

rates or, conversely, without selection of the proper converter size for a given production, inefficient operation may result, not only from too high a velocity with the aforementioned yield loss but also from too low a velocity, which can produce gas channels along the walls of the converter.

The catalyst generally used for ammonia oxidation is a platinum–rhodium alloy containing about 10 percent rhodium. An 80 mesh gauze made of wire 0.003 inch in diameter is used. During the life of the catalyst, it becomes polycrystalline in appearance. There is also some weight loss, the effect of which is more pronounced at the higher converter temperatures employed in the pressure process. With the atmospheric processes at lower temperature, the platinum loss is about 45 mg/metric ton of acid produced. Precious metal recovery steps are required for operation at 120 lb pressure. This usually involves periodic cleaning of the converter and heat exchange equipment, and the use of a filter in the gas stream beyond these pieces of equipment.

A summary of catalyst developments was given in a British report.[53] The materials examined as a catalyst were platinum, platinum–iridium alloys, and platinum–rhodium (various compositions; see Fig. 28.13). Other alloys of platinum were examined, such as those containing palladium, ruthenium, rhenium, and tungsten. Today practically all ammonia oxidation processes in the United States use a combination of platinum–rhodium–palladium with the palladium present to replace a portion of the more expensive rhodium. Loss of activity results from contaminants such as iron, calcium, MoS_2 (lubricants), and phosphorus in the gas streams.

Fig. 28.13. Effect of rhodium content on catalyst efficiency.[53]

Several filter systems to recover catalyst that would otherwise be lost have been investigated. These include filters of glass wool, asbestos fibers, Raschig rings, and marble chips (3–5 mm dia.). Shortcomings of these recovery systems have been overcome by a system that uses palladium alloy gauzes after the platinum gauze. The gauze is placed as close to the production catalyst as possible, the idea being to recover the catalyst components while it is in the vapor form. Palladium, by itself, is reported to be the most effective "getter," but when used alone, the palladium becomes too brittle. Thus, a palladium alloy generally is used. The total recovery, S, for a number of gauzes can be estimated by the following:

$$S = 1 - \left(\frac{a}{100}\right)^n \times 100$$

where:

a = % recovery by the first gauze
n = number of gauzes

Normally a recovery of 55 to 70 percent is economical.

Base metal catalysts have been investigated in fixed-bed and fluidized-bed systems but have not proved to be effective. The following materials have been tried: Fe_2O_3–Mn_2O_3–Bi_2O_3 (Braun oxide); CoO–Bi_2O_3; and CoO with Al_2O_3, thorium, cerium, zinc, and cadmium.

Absorption of nitric oxide takes place in a sieve or bubble-cap tower into which air is added to oxidize the nitric oxide to nitrogen dioxide. The nitrogen dioxide must be absorbed in water to liberate nitric acid and nitric oxide. There are two equilibria, as follows:

$$N_2O_4 \rightleftarrows 2NO_2 \rightleftarrows 2NO + O_2$$

The first of these two equilibria is established relatively quickly; the second more slowly. At 150°C the nitrogen tetroxide is almost completely decomposed into nitrogen dioxide. At 500°C the equilibrium lies at 25 percent nitrogen dioxide and 75 percent nitric oxide. The formation of nitrogen dioxide from nitric oxide and oxygen is relatively slow at atmospheric pressure, but proceeds quite rapidly at higher pressure. Because this reaction has a negative temperature coefficient, the absorber is operated at the lowest temperature obtainable with available cooling water, in the temperature range of 10 to 40°C.

U.S. Patent 4,957,720[54] covers an improvement to the nitric acid process in which process heat is better utilized thereby improving the economics. Recovery of heat at relatively low temperature from the NO_x gases is accomplished.

Concentration of Nitric Acid. The most widely used method of concentrating nitric acid consists of mixing the 60 percent nitric acid with strong (93%) sulfuric acid, and then passing the mixture through a distillation system from which concentrated (95–98%) nitric acid is taken overhead and denitrated. Residual sulfuric acid containing approximately 70 percent H_2SO_4 is obtained, and the dilute residual sulfuric acid may be concentrated for reuse.

A process that employs magnesium nitrate as a dehydrating agent is illustrated in Fig. 28.14. An aqueous solution containing 72 percent magnesium nitrate is fed to the middle of the tower, which also receives the 60 percent nitric acid feed. Pure nitric acid of 95 to 98 percent strength is taken overhead. A solution containing about 55 percent magnesium nitrate is removed from the base; and after reconcentration to 72 percent, the magnesium nitrate is recirculated to the nitric acid tower.

The Pintsch Bamag process[55] concentrates nitric acid without dehydrating agents by using the difference in the composition of the nitric acid–water constant-boiling mixtures at different pressures. A two-column distillation system is employed.

Uhde has developed another process for producing highly concentrated nitric acid (98–99%) without using dehydrating agents. The NO from the ammonia burner is converted to NO_2 by means of highly concentrated HNO_3. The NO_2 is cooled with brine, and after removal of the residual water, it is physically absorbed in chilled, highly concentrated nitric acid. Then it is separated from the concentrated nitric acid in a distillation column, and liquefied by refrigeration whereby the NO_2 is dimerized in a $N_2O_4/HNO_3/H_2O$ mixture, which reacts with supplemental oxygen at a pressure of about 720 psi to form highly concentrated nitric acid.

Fig. 28.14. Magnesium nitrate process for nitric acid concentration. (Hydrocarbon Processing, ***58**, No. 11, 209, Nov. 1979, copyright 1979 by the Gulf Publishing Company. By permission.*)

The acid leaving the reactor contains about 20 percent dissolved N_2O_4, which is separated in a distillation column and returned to the reactor cycle. The product is bleached and sent to storage, with a portion recycled to the absorber. (See Fig. 28.15.) Figure 28.16 gives a view of a Uhde plant for production of concentrated nitric acid.

Typical consumption figures for a medium-sized plant (150 tons/day), per ton of HNO_3 (100%), are as follows:

NH_3	282 kg
O_2	125 m^3 @ NPT
Electricity	285 kWh
Cooling water	200 m^3
Excess steam	0.6 t
NO_x in waste gas	200 ppm

1. Air filter
2. Ammonia-air-mixer
3. Mixed gas filter
4. Burner with waste-heat boiler
5. Gas cooler I
6. NO blower
7. Gas cooler II
8. Oxidation column
9. Circulating acid tank
10. Acid cooler
11. Final oxidator
12. Brine gas cooler
13. Absorption column
14. Tail gas scrubber
15. Head tank for raw acid
16. Raw acid cooler
17. Liquefier
18. Agitator vessel for raw mixture
19. Reactor vessel
20. Bleaching column
21. Final acid cooler

Fig. **28.15**. Concentrated nitric acid by Uhde. (*Courtesy of Uhde.*)

Fig. 28.16. Plant using Uhde process for direct production of concentrated nitric acid. (*Courtesy of Uhde.*)

Purification. Chemically pure nitric acid is required for some applications and as a reagent for analytical measurements. The commercial grade may contain impurities, such as arsenic, sulfur, and phosphorus.

In one purification process (U.S. Patent 3,202,481)[56] potassium or sodium permanganate is used to oxidize the impurities. Other oxidizing agents also can be used, such as persulfates or chlorates. Molybdenum trioxide is used to form a complex with the oxidized phosphorus impurities. The nitric acid is contacted with the oxidizing agent and the molybdenum compound at the boiling point of the acid, and the acid is continuously removed by distillation. The oxidized impurities remain in the residue. As a result, the purified acid will contain less than 0.5 ppb by weight of arsenic and phosphorus.

Stabilizers. Over a period of time, concentrated nitric acids tend to decompose according to the following equation:

$$4HNO_3 \rightarrow 4NO_2 + 2H_2O + O_2$$

As a result, pressures will build up in storage vessels. Also nitric acid is very corrosive. Therefore, some stabilizers and/or corrosion inhibitors are used.

Corrosion of aluminum by red fuming nitric acid is reduced by adding 4 percent by weight of hydrogen fluoride, 52 percent HF. (See U.S. Patent 2,760,845.)[57]

Decomposition of the concentrated acid is reduced by such substances as quaternary ammonium compounds, organic sulfones, inorganic persulfates, and organic sulfonium compounds.[58]

Pollution Abatement. The tail gas from ammonia-oxidation processes may contain sufficient nitrogen oxides to result in serious atmospheric pollution (smog) or serious corrosion of power-recovery equipment, if such is used, unless some means for removal of the oxide is employed. The oxides usually are reduced with natural gas, hydrogen, or a carbon monoxide–containing gas, or they can be absorbed by alkaline liquors (soda, calcium, or magnesium) or ammonium sulfite–bisulfite.[58]

As an example of the reduction procedure, a gas containing the materials listed below was treated with 100 scf of natural gas and 1000 scf of air over a nickel catalyst on alumina at 90 psig, 1150°F (1600°F exit), and with a space velocity of 20,000 cu ft of exit gas/cu ft of catalyst.

Gas	%
N_2	97.2
O_2	2.0
H_2O	0.6
$NO + NO_2$	0.2

The effluent gas contained about 75 ppm NO–NO_2 mixture; 0.1 percent CO; 0.3 percent H_2; and 1.5 percent CO_2.

With methane as the reducing gas, materials that are particularly effective as catalysts are rhodium, palladium, platinum, and ruthenium. Alumina ceramic honeycomb material can be used as a base.

These catalysts usually are used at about 0.5 percent by weight on Al_2O_3. With these materials, the initial reaction temperature is from 700°F to 1000°F, depending upon the catalyst and oxygen concentration.

Harding and others[59] developed an absorption–desorption process using zeolites. It is stated that this process will result in the

recovery of the oxides and add up to 4 to 5 tons/day of capacity to a 300-ton/day plant. Exit NO_x concentrations of 0.001 percent were obtained for 3 hours and 0.02 percent after 3 hours when 2.4 scfm tail gas (0.18 percent nitric oxides) was fed over 4.9 lb of zeolite. The columns were regenerated with steam at 300 to 350°F, or by hot air.

Another method of abating tail gas from ammonia-oxidation processes that normally do not exceed 2000 ppm nitrogen oxides is selective catalytic reduction (SCR). The term "selective" is used because the ammonia fuel only reacts with the nitrogen oxides and not with the oxygen first as in other abatement systems. Catalyst materials include vanadia/titania and iron–chromium compounds.

Production. Besides the military, major producers are: Du Pont, Arcadian, Chevron, Columbia Nitrogen, and Mississippi Chemical. Production is declining slowly in the United States, primarily because of a shift away from nitrate fertilizers, a reduced need for explosives, and some increase in imports. Domestic production capacity peaked in the early 1980s at near 11,000,000 short tons/year. Since that time, at least 13 companies have divested themselves or closed down nitric acid production capacity. By 1987 production capacity had declined to 9,800,000 short tons. Demand between 1981 and 1987 declined from 10,000,000 to 7,400,000 short tons, and it was less than 6,700,000 short tons in 1991.[60] In other parts of the world, Uhde has built plants with capacities up to 1800 tons/day (Norsk Hydro, Oslo-Norway), and Grande Pariosse has built a plant with capacity up to 1250 tons/day (Sluiskil, Holland). A Du Pont plant in Texas was built in 1977 with a capacity of over 1000 tons/day.

Uses of Nitric Acid

The primary use of nitric acid is in the manufacture of ammonium nitrate for the fertilizer and explosives industries. Approximately 64 percent of the total production goes into fertilizers. Solution fertilizers, many of which contain nitrates, are growing in importance. Nitric acid is used in some cases for the acidulation of phosphate rock to produce mixed fertilizers. (See Chapter 11.)

The second most important use, accounting for about 16 percent of the nitric acid produced, is for textile fibers, plastics, and elastomers. Approximately 1.2 million tons are used in this market with 700 thousand of them used in adipic acid, an intermediate for nylon.

About 15 percent of the nitric acid produced is consumed in industrial explosives (see Chapter 30). Approximately 1.1 million tons/year of nitric acid is used for explosives, which, in turn, are used in coal mining, quarrying, metal and nonmetal mining, heavy construction, and petroleum exploration. Fertilizer-grade ammonium nitrate is used as an explosive or blasting agent. It is mixed with oil and set off by priming with a high explosive.

Nitric acid has a number of other industrial applications. It is used in pickling stainless steel, in steel refining, and in the manufacture of dyes, plastics, and synthetic fibers. Most of the methods used for the recovery of uranium, such as ion exchange and solvent extraction, use nitric acid. Thus, with the expansion of the nuclear industry, nitric acid consumption may increase. The list price for nitric acid in 1991 was $185.00/ton for 60 percent in tanks C.I. works, 100 percent basis.[46]

Licensers for nitric acid processes in 1990 were:[47]

CECP	Montedison
Chemetics	Polimex
Corning Process Systems	Quimigal
Davy McKee	Rhone-Poulenc
Deutsche Engineering	Stamicarbon
Espindesa	Sumitomo Chemical
Grand Paroisse	Uhde
Hercules	Weatherby

AMMONIUM NITRATE

Ammonium nitrate was first prepared in the sixteenth century. Its early industrial development was primarily for use in explosives, but beginning in about 1940 its use as a fertilizer

developed rapidly. Ammonium nitrate is the most important ammonium compound, but during the 1980s its production declined at an average of 2.8 percent/year in the United States, to 7,611,000 short tons in 1990.[2,3]

This decline has been due to a decline in the U.S. consumption of ammonium nitrate as fertilizer over recent years. During this period production fell to a low of 6.5 million short tons in 1986, but it recovered with the help of declining imports and rising exports. Large amounts of ammonium nitrate still are consumed for explosives and in the manufacture of nitrous oxide. The process for manufacturing ammonium nitrate consists of neutralizing ammonia with nitric acid under controlled conditions:

$$HNO_{3(l)} + NH_{3(g)} \rightarrow NH_4NO_{3(aq)} + 26 \text{ kcal/g mol evolved}$$

Water is evaporated, and the anhydrous, or nearly anhydrous, solution is prilled or spherodized to a solid product. Figure 28.17 shows the C&I/Girdler, Inc. process with either the prilling tower or the spherodizer. Anhydrous ammonia and 50 to 60 percent nitric acid are fed to the neutralizer, where the water is evaporated by the heat of reaction (550–620 Btu/lb of ammonium nitrate, depending on the acid strength). Final concentration is done in a falling film evaporator in which a stream of air helps to carry off the water vapor. The air stream from the evaporator is mixed with that from the spherodizer for cleanup in a wet scrubber.

A C&I/Girdler additive or nucleating agent for the prilling process is added to the melt before prilling or spherodizing. This produces unstressed crystals and stabilizes the prills against temperature cycling through crystal transition phases. It reduces prill breakage and the amount of coating agent required. The product from the prill tower or spherodizer is sent to a final cooler.

In the spherodizer granulation process, the molten ammonium nitrate is sprayed onto the rolling bed of solid particles in the drum, forming granules, which are cooled and screened. Advantages of spherodizing include the following: there is variation in product size, atmospheric pollution from the prill tower is eliminated, and a harder product is made. A yield of over 99.5 percent from ammonia and nitric acid is claimed.

To prill ammonium nitrate, a solution is evaporated in a vacuum or falling-film

Fig. 28.17. Ammonium nitrate manufacture per C&I Girdler, Inc. (Hydrocarbon Processing, *58, No. 11, 1979, copyright 1979 by Gulf Publishing Company and used by permission of the copyright owner.)*

evaporator until it is almost free of water. The concentration of ammonium nitrate will vary from 95 to 99.95 percent. The prilling tower itself is about 50 to 100 feet tall, depending on the process being used. The concentrated ammonium nitrate solution is fed through spray-formers, which break up the stream into small droplets that solidify during free fall in the tower. In some instances limestone is mixed with the ammonium nitrate before prilling; in other cases, the prills are coated with limestone dust. A mixture of limestone and ammonium nitrate is known as "nitro chalk" in the fertilizer industry. Plants based on 5 percent moisture in the feed to the prilling tower usually require a drying step before the material can be bagged.

There are five different crystalline forms of ammonium nitrate.[48] One is stable below −18°C, another at −18 to 32°C, a third at 32 to 84°C, a fourth at 84 to 125°C, and a fifth at 125 to 169.6°C, the melting point. The molecular weight is 80.05, the nitrogen content 35 percent, and the solubility in water 118.3 g/100 g at 0°C and 576 g/100 g at 80°C.

Ammonium nitrate decomposes in two ways. There can be controlled decomposition with careful heating, which is the commercial process for producing nitrous oxide:

$$NH_4NO_3 \xrightarrow{200\text{–}260^\circ C} N_2O_{(g)} + 2H_2O_{(g)} + 6720 \text{ cal/g mol}$$

There also can be explosive decomposition by heating or by explosive shock, which is the basis for its use as an explosive:

$$2NH_4NO_3 \rightarrow 2N_2 + 4H_2O + O_2 + 28{,}200 \text{ cal/g mol}$$

The oxygen that is released can be used to more than triple the explosive effects by mixing an organic, usually fuel oil, with the ammonium nitrate. This mixture, which goes by the acronym ANFO, has revolutionized the explosives business over the past 40 years because ammonium nitrate is so much cheaper and so much safer to handle than dynamite.

$$3NH_4NO_3 + (CH_2)_n \rightarrow 3N_2 + 7H_2O + CO_2 + 102{,}500 \text{ cal/g mol}$$

It is important to mention the hazards involved in handling ammonium nitrate and the precautions that should be taken. Large masses of concentrated solutions at high temperatures should be avoided; in particular, one should avoid contamination with decomposition catalysts such as chlorides and organic materials. Under process conditions, one should keep the melt above a pH of 4.5, and keep the neutralizer well ventilated.

One of the most famous disasters in U.S. history occurred in 1947 when a ship loaded with ammonium nitrate fertilizer exploded at the dock in Texas City, Texas. The Texas City disaster was dramatic proof that organic material and ammonium nitrate decompose with explosive violence. That explosion took almost 600 lives, injured 3500, and destroyed 33 million dollars worth of property. Ammonium nitrate solutions also can explode with disastrous results.

Stafford, Samuels, and Croysdale[61] gave a comprehensive review of ammonium nitrate plant safety. They discussed how coating ammonium nitrate affects its explosibility. For instance, 2.5 percent by weight of diatomaceous earth applied as a coating on ammonium nitrate reduces its detonation sensitivity, and 2.5 percent by weight of finely divided sulfur increases the detonation sensitivity. At a 5 percent concentration, urea decreases the thermal stability of ammonium nitrate.

Some precautions that have been suggested in the manufacture of ammonium nitrate are as follows:[62] (1) operation should be conducted below 250°F; (2) when the pH is less than 4.0 for a prolonged period of time, type 304L stainless steel should be used; and (3) when the pH is over 4.0 for solids and over 6.0 for solutions and the temperature is less than 240°F, aluminum may be used.

By far the most important use for ammonium nitrate, accounting for over 80 percent of the ammonium nitrate produced, is as fertilizer. Containing 34 percent nitrogen in the form of nitrate and ammonia, both of which are rather easily assimilated by plants, nitrate continues to find wide use for this purpose.[129] Before about 1940, the use of ammonium

nitrate as a fertilizer was hampered by the hygroscopic nature of the crystalline material. About that time it was discovered that with certain additives and with proper handling in terms of temperature and moisture exposure, the material could be kept relatively free-flowing. In recent years, formulations and mixtures of ammonium nitrate with other chemicals have continued to be developed, either to improve its fertilizing properties, to make its handling easier, or to make the material safe from explosions. Large-scale production facilities for UAN, which is a urea–ammonium nitrate solution, have been built for fertilizer use.[44]

Included in these improvements are additions of stabilizing proprietary agents that prevent breakdown of the crystals on storage. Solutions also have been developed that are blends with other fertilizers such as urea to make handling easier and to provide improved fertilizing properties.

Uhde GmbH licenses an improved control system for ammonium nitrate plants that gives up to a 99.8 percent yield.[63]

The list price for ammonium nitrate in 1991 was $126/ton in bulk, delivered.

Licensers for ammonium nitrate plants are:

Deutsche Engineering	Montedison
Espindesa	Polimex-Cekop
Grand Paroisse	Quimigal
Hoechst	Stamicarbon
Hydro Agri	Sumitomo Chemical
Incro	Uhde
Kaltenbach-Thuring	Weatherly

UREA

Urea, also named carbamide or carbonyldiamide, was first synthesized by Wohler in 1828 from ammonia and cyanic acid. It is a colorless crystalline material, soluble in water and in alcohol but not in ether. Urea contains 46 percent nitrogen, the most of any ordinarily solid fertilizer material. Ammonium nitrate contains 34 percent nitrogen. Urea is not flammable, but will melt and decompose in a fire to give off ammonia. Table 28.15 lists its physical properties.

TABLE 28.15 Physical Properties of Urea

Chemical formula	NH_2CONH_2
Molecular weight	60.06
Freezing point, °C	133
Boiling point, °C	decomposes
Specific gravity	1.34
Bulk density	0.70
Heat of combustion	-91.02×10^5 J/kg
Heat of solution	-2.52×10^5 J/kg

Because urea is converted into ammonia and then into nitrates in the soil, it makes a very concentrated form of nitrogen fertilizer. It also is used in resin and plastics manufacture and in the synthesis of organic materials.

The production of urea worldwide has continued to increase, with a growth rate of near 7 percent in the 1980s. Production in 1988 was estimated to have exceeded 155 billion lb.[130]

Urea can be produced by the hydrolysis of cyanamide according to the following equation:

$$\underset{\text{Cyanamide}}{CNNH_2} + H_2O \rightarrow \underset{\text{Urea}}{CO(NH_2)_2}$$

At high pressure (33–53 atm) urea also can be formed by heating ammonium carbonate. Equilibrium is obtained at 130 to 150°C with a 30 to 45 percent yield. A process was described by Raymond Fanz and Fred Applegath of Lion Oil Company that produces urea from ammonia, carbon monoxide, and sulfur (dissolved in methanol).[64] The process operates at 100°C and 20 atmospheres and produces H_2S along with the urea. The H_2S is oxidized to sulfur for recycle.

The currently favored method of manufacture is to combine liquid ammonia with carbon dioxide under pressure to form ammonium carbamate, which then is dehydrated into urea in a water solution. The unreacted carbon dioxide and ammonia are recovered and recycled to the synthesis

operation

$$2NH_3 + CO_2 \rightleftarrows \underset{\text{Ammonium carbamate}}{NH_2COONH_4} \text{ (exothermic)}$$

$$NH_2COONH_4 + \text{heat} \rightarrow \underset{\text{Urea}}{NH_2CONH_2} + H_2O \quad \text{(endothermic)}$$

There now are three prevailing process designs for urea manufacture, licensed by Stamicarbon, Mitsui Toatsu, and Snamprogetti. The main differences between these processes are in the methods used to handle the converter effluent, to decompose the carbamate and carbonate, to recover the urea, and to recover the unreacted ammonia and carbon dioxide for recycle, with the objective being minimum expenditure of energy and a maximum recovery of heat. In some processes, a liquid is used to recycle a solution of carbamate, carbon dioxide, ammonia, and water. In others, the amount of water recycled is minimized, and only carbon dioxide and ammonia are recycled. In the older plants, or once-through plants, the off-gases are used as feed to ammonium nitrate or ammonium sulfate plants.

The Mitsui Toatsu Chemical (MTC) process was the prevailing process into the early 1970s and generally serves as a base for comparing new processes. The urea reactor operates at high pressure (250 atm). The reactor liquid effluent is let down in stages, and steam is used to decompose carbamate and strip the CO_2 and NH_3 in each stage. The CO_2 and NH_3 are absorbed in water for recycle to the reactor.

Although the CO_2 and NH_3 stripping processes gained favor, MTC-based plants still were being built in the early 1980s, and MTC continues to improve and optimize its process.

The Snamprogetti process is based on using ammonia to strip CO_2 and NH_3 out of the urea reactor effluent. One of the impurities that affects the performance of urea as a fertilizer is biuret. To avoid a yield loss and minimize product contamination, it is important to minimize biuret formation. The concentration of biuret in urea used for foliar feeding must be kept low to avoid mottling of the leaves. Consequently, a foliar grade of urea is produced, containing less than 0.035 percent biuret. Its formation is as follows:

$$\underset{\text{Urea}}{CO(NH_2)_2} \rightleftarrows HNCO + NH_3$$

$$CO(NH_2)_2 + \underset{\text{Cyanic acid}}{HNCO} \rightleftarrows \underset{\text{Biuret}}{NH_2CONHCONH_2}$$

The rate of conversion reportedly is increased by basic compounds and decreased by acidic compounds.

A low-biuret product can be produced by recycling the biuret-containing effluents from the purification zone back to the converter. See U.S. Patent 3,232,984.[65] The stripping NH_3 actually is fed to the reactor as excess, which inhibits the formation of biuret. Snamprogetti also claim lower maintenance costs, and many new plants based on this process were installed in the early 1980s.

Mitsui Toatsu Chemicals and Snamprogetti offer plants with capacities of up to 1800 metric tons/day. A variety of modifications are available with these processes, the latest of which incorporate energy-saving developments. The Snamprogetti process is shown in Fig. 28.18.[66]

The Stamicarbon process is based upon the use of carbon dioxide to strip ammonia from the reactor effluent countercurrently. This process has become the most popular with new plants. As the carbon dioxide removes the ammonia from the solution, the carbamate decomposes, leaving a minimum amount in the effluent. The solution leaves the stripper at about 150 to 180°C. The stripped gases flow to the reactor along with ammonia in an amount equivalent to the amount of carbon dioxide added for stripping. Most of the off-gas from the reactor is condensed, and inert gases are bled from the system before the condensate is returned to the base of the reactor. This process is shown in Fig. 28.19.

For a Stamicarbon total recycle plant with CO_2 stripping to produce fertilizer grade prills, typical consumptions per metric ton of

product are shown below:

	Stamicarbon	*Snamprogetti*	*Mitsui Toatsu*
kg CO_2	735	735	735
kg NH_3	570	570	570
kWh power	21	38	17
kg steam (net) (5 atmg)	800 (55 atm)	1290 (37 atm)	1100 (95 atm)
m^3 cooling water ($T = 11°C$)	80	65	90

Fig. 28.18. Snamprogetti NH_3 stripping process. (*Courtesy Bechtel Corporation*).

Fig. 28.19. Stamicarbon process for urea. (Hydrocarbon Processing, *58, No. 11, 1979, copyright by Gulf Publishing Company and used by permission.*)

The Allied Chemical—C.P.I. process, which was developed by T. O. Wentworth,[67] is based on recycle of NH_3 that is separated from the reactor effluent gases by scrubbing with monoethanolamine. The carbon dioxide is absorbed in the amine, and the gaseous ammonia is recycled to the reactor. (See Fig. 28.20.) The reactor effluent first is decomposed; then the urea solution is removed and concentrated. A specially designed reactor lined with zirconium is needed. Carbon dioxide conversion is relatively high—80 to 85 percent per pass.[68]

The Chemico "Thermo-Urea" process[70] employs centrifugal compressors to recycle the carbon dioxide, ammonia, and water vapor to the reactor (Fig. 28.21). The reactor effluent is passed into a separator from which the hot gases pass to the compressor, and the liquid effluent enters a carbamate decomposer

Fig. 28.20. Allied Chemical-C.P.I. process for urea. (*U.S. Patent 3,107,149; Noyes; U.S. Patent 3,236,888.*)[67–69]

Fig. 28.21. Chemico thermo-urea process. (*U.S. Patent 3,200,148.*)[70]

where the remaining gases are removed and also sent to the compressor. The liquid effluent from the first decomposer enters a low-pressure decomposer, which removes the last traces of carbon dioxide, ammonia, and inert gas. The ammonia and carbon dioxide are absorbed in water for recycle to the compressor after vaporization, and the inert gas is removed from the top of the absorber. Aqueous urea is removed from the low-pressure decomposer. An improved process is described in U.S. Patent 3,301,897.[71]

Anhydrous urea can be produced directly, according to J. S. Mackay,[72] by synthesizing the urea at a temperature above its melting point and simultaneously keeping the partial pressure of the water below the condensation point. A pressure above 1000 psig and a temperature above 200°C are used. The preferred conditions are 2000 psig and 225°C. Nitrogen, but preferably ammonia, can be used to reduce the partial pressure of the water vapor.

$$10NH_3 + CO_2 \rightarrow \underset{\text{Urea}}{(NH_2)_2CO} + H_2O + 8NH_3$$

Product Formulation

The urea–water solution produced by the various processes must be processed further to obtain the final products, in the form of crystals, flakes, prills, or granular material. By far the largest amount of urea is distributed as a fertilizer in the form of prills, which are produced by removing water from the solution and dropping the molten urea down through spray formers in a tower countercurrent to a stream of air. The droplets of molten urea, which solidify before hitting the base of the tower, are cooled, and then are bagged or sold in bulk.

Crystallization of urea from the urea–water solution produces a relatively pure product. The biuret content is lower than that in the standard prilled product because the biuret in the urea–water solution is removed with the mother liquor in the crystallization step. Prills with low biuret content can be produced by starting with crystallized urea.

Granular material is produced by mixing concentrated urea solution into an agitated mass of dry particles. The solution is at a temperature a little higher than its set point. Agitation is continued until granules are obtained; then they are dried at a temperature of no more than 120°C. The usual ratio of urea solution to dry particles is 0.4:1 to 0.7:1.

Urea absorbs moisture, which causes caking. The critical humidities (% relative humidity above which prills will absorb moisture) are 80 percent at 10°C, 76 percent at 25°C, and 69 percent at 40°C. The techniques used to reduce caking are: (1) coating, (2) mixing with additives, and (3) thermal treatment.

Coatings that are reported to be effective are sugar, formaldehyde, and acetaldehyde. Additives such as ammonium sulfate, acetylene–diurea, magnesium carbonate, and substituted ureas (e.g., *p*-toluidine) may be used. In the thermal treatment, the prills are passed through a zone of temperature higher than the melting point of urea.[68]

Uses of Urea

The principal use for urea is as fertilizer, with over 80 percent of all production being consumed for this purpose. Urea is used as a solid fertilizer and also in solutions in combination with ammonia and ammonium nitrate. The latter commercial product is known as UAN. Continued growth is expected because the solution can be handled as if it were water, in contrast to the handling required for anhydrous ammonia.

In 1982, 8,620,000 short tons of urea were produced in the United States. Since that time U.S. production has declined because of reduced exports and increased imports.

In the world market, urea is the leading nitrogen fertilizer material.[73] Its high nitrogen analysis saves transportation costs; also, urea can be used on almost any crop, and its production costs have been lowered by the large single-train ammonia and urea plants. All indications are that the U.S. fertilizer market is moving toward the greater use of fluids. When combined with superphosphate, urea produces a fertilizer known

as "Phosphazote." A urea salt with calcium nitrate also is recommended as a fertilizer. In addition, urea can be combined with formaldehyde to produce a slowly soluble compound that will supply nitrogen to the soil during the entire growing season. Worldwide, urea continues to grow at 10 to 15 percent a year.

Industrially, when combined with formaldehyde, urea gives resins that can be machined. It is used as an ingredient in softeners for cellulose, cellophane, and wood, and to retard end-checking of boards during drying. Added to glue, gelatin, or starch, it reduces the viscosity and permits the use of higher concentrations of the active material. In the petroleum industry, urea is used to separate straight-chain hydrocarbons by forming crystalline complexes between the urea and a hydrocarbon with a long straight carbon chain. Urea is used in the preparation of barbituric acid, caffeine, ethylurea, hydrazine, melamine, guanidine, and sulfamic acid. It also finds use in the manufacture of medicinals, in some cosmetic applications and deodorants, and in skin creams to improve the texture. With hydrogen peroxide, urea forms a crystalline additive $CO(NO_2)_2 \cdot H_2O_2$ that is employed as a disinfectant and an oxidizing agent.

Urea–formaldehyde, urea–furfural–formaldehyde, and urea–resorcinal mixtures form thermosetting resins that are used in molding applications such as radio and TV cabinets, in the treatment of paper and fabrics, and as adhesives for bonding wood. The resins are nearly colorless but opaque, and may be pigmented as desired. They improve crush and crease resistance in textiles, and impart wet strength to paper and paper coatings. A small amount of acid, or acid-generating catalysts, may be added to make the resin set.

Urea has found acceptance as a protein substitute for ruminant animals (cattle, goats, and sheep). In this application, it can efficiently replace about one-third of the natural protein in ruminant feeds, but care in formulating must be exercised to optimize the concentration. Urea is toxic in large doses. It is reported[60] in a recent patent that urea–hydrocarbon solids adducts are more palatable than urea alone, making the product more readily accepted by the animals. The hydrocarbons used are alkanes, such as *n*-hexane to *n*-tetradecane, and olefins, such as *n*-hexene and *n*-octadecene.

The following table shows the advantage of a urea adduct over urea:

Feed Additive	*Nitrogen Retained (g/day)*
Urea	1.72
n-Tetradecane–urea adduct	2.16
1-Tetradecene–urea adduct	2.22

The price for urea has risen modestly in recent years:

	Price per short ton[46,60]
1978	$125
1983	135
1991	160

Licensers for urea plants are as follows:[47]

Bechtel	Quimigal
Hydro AgriEurope	Snamprogetti
Inco (Cros)	Stamicarbon
Kaltenback–Thuring	Toyo Engineering
Montedison	Weatherly

MELAMINE

Melamine (1,3,5-triazine-2,4,5-triamine; $C_3H_6N_6$) was first prepared by Liebig in 1834. It was not until 100 years later, in the 1930s, that it received commercial attention. Since then it has developed steadily until today world capacity exceeds one billion pounds. Melamine or cyanuramide, $C_3N_3(NH_2)_3$, is a monoclinic fine white crystalline powder. Its insolubility in most organic solvents but appreciable solubility in water are two important properties. It is only mildly toxic.

Three raw materials, urea, hydrogen cyanide, and dicyandiamide, a derivative of calcium cyanide, can be converted to maleamine. The route using dicyanamide was the original process, developed in the 1930s, but today practically all melamine worldwide is produced

from urea. The HCN route never was commercialized.

Melamine is produced by heating urea and passing the resulting mixture of isocyanic acid and ammonia over a suitable catalyst, quenching the products with water or aqueous mother liquor, and working them up by filtration, centrifuging, or crystallization.

The reaction takes place in two stages. The first reaction is:

$$\underset{\text{Urea}}{CO(NH_2)_2} \rightarrow \underset{\text{Cyanic acid}}{HOCN} + NH_3$$

This reaction is endothermic, requiring 800 kcal/kg of urea. The second reaction is:

Melamine

This reaction is exothermic, liberating 1100 kcal/kg of melamine. The overall reaction is endothermic.

The process can be carried out in the liquid phase by applying high pressure (100–150 atm) and without a catalyst. This process was developed in the early 1960s. Synthesis from urea also can be carried out in the vapor phase at lower pressure (< 10 atm) using alumina catalyst.

The ammonia and the carbon dioxide of reaction are recycled to produce more urea, substantially improving the process economics.

The O.S.W. process is shown in Fig. 28.22.[74] Urea, along with recycle ammonia, is fed into a fluidized-bed reactor containing inert solids, where the temperature is raised to 350°C within a few seconds. The resulting mixture of isocyanic acid and ammonia then is fed into a reactor containing a solid catalyst, such as silica gel, alumina, or boron phosphate. Products from the isocyanic acid reactor are quenched with water, the resulting slurry is centrifuged, and the mother liquor is sent to recovery or a urea plant.

Off-gases from the quench system contain ammonia, carbon dioxide, and water vapor. The CO_2 is removed as ammonium carbonate, and the ammonia is returned to the urea reactor.

The design of the urea reactor is critical. The wall temperature must not be too high, or corrosion is a problem.[75] This can be avoided by using stationary layers of catalyst along the inner wall, held in place by corrosion-resistant flights secured to the periphery at suitable intervals. With this

Fig. 28.22. Melamine synthesis-O.S.W. (Hydrocarbon Processing, *48, No. 11, 209, Nov. 1969.*[74] *Copyright 1969 by Gulf Publishing and used by permission of the copyright owners.*)

method, the reactor wall can be made of steel containing 0.16 percent carbon, 0.35 percent silicon, 0.40 to 2.00 percent manganese, not more than 0.5 percent phosphorus, and not more than 0.05 percent sulfur.

The formation of melamine from isocyanic acid takes place at 400 to 430°C. The heat of reaction is absorbed as sensible heat by raising the entering vapors from 320°C to 430°C. Melam ($C_6H_9N_{11}$) and melem ($C_6H_6N_{10}$), condensation products of melamine, are by-products that form by splitting off ammonia when the temperature is too high. Melamine will hydrolyze if allowed to remain too long with water at high temperature, the resultant products being ammeline $(CH)_3(NH_2)_2OH$ and ammelide $(CH)_3(NH_2)(OH)_2$.

Melamine is milled to produce a fine particle size, 5 to 10 microns.

Methods of purifying crude melamine are being studied[76] because recrystallization is required to obtain a pure product. In the reference cited, crude melamine is dissolved in aqueous ammonia, from which it is crystallized. The obtainable yields are reported to be 10 percent higher than for conventional processes.

Melamine is used in the production of resins, primarily melamine–formaldehyde (MF) resins. Surface coatings account for about 36 percent of resin use, and that use is growing. Major uses of melamine are as follows:

Surface coatings	36%
Laminates	26
Paper coatings	14
Plastic dishware	5
Textile coatings	5
Wood adhesives	5
Other	balance

U.S. production is estimated at about 150 million lb/year.[77]

There are two producers in the United States. American Melamine Industries (AMEL) is a joint venture between American Cyanamid and Dutch State Mines (DSM). The 120 million lb/year plant is located on American Cyanamid's site at Fortier, Louisiana. Melamine Chemicals has a 105 million lb/year plant at Donaldsonville, Louisiana. As much as 20 million lb/year is imported, mostly from DSM in Europe.

The price for melamine in 1991 in bags, f.o.b. works, was 66 cents/lb.

Licensers for melamine plants are Allied-Signal, Granuit, and Stamicarbon.[47]

ALIPHATIC AMINES

Amines are derivatives of ammonia in which the hydrogen atoms in the ammonia have been displaced by monovalent hydrocarbon radicals. Depending on the number of hydrogen atoms displaced, amines are classified as primary (RNH_2), secondary (R_2NH), and tertiary (R_3N). According to the type of hydrocarbon radicals that substitute for the hydrogen atoms, amines are categorized further as aliphatic (saturated or unsaturated), aromatic, heterocyclic, alicyclic, or any combination of these. Amines are distinguished from imines, in which two hydrogen atoms in an ammonia molecule are displaced by bivalent hydrocarbon radicals ($R{=}NH$); and they are distinguished from nitriles or cyanides, in which all the hydrogen atoms in ammonia are displaced by a trivalent hydrocarbon radical ($RC{\equiv}N$).

Although the amine family consists of a multitude of compounds, the scope of this section is confined to the following commercial aliphatic amines produced in large volume:

- Methylamines
- Ethylenediamine

In general, most aliphatic amines are toxic and corrosive to skin. Low molecular weight amines are volatile and may cause bronchitis and even suffocation if inhaled. One should wear full protective equipment when handling the material. The first aid treatment upon superficial contact with amines is to flush the exposed area with large quantities of water.

Methylamines

Early commercial use of methylamines was in the leather industry as a depilatory agent.

Since the 1950s, amines have been developed into major industrial chemicals, with end uses ranging from military rocket fuels to building blocks for pharmaceuticals, solvents, and agricultural chemicals.

Methylamines are colorless liquids that are volatile under atmospheric conditions. The amines have threshold odor limits of less than 10 ppm; at low concentration they smell fishy, whereas at high concentration they smell like ammonia. Nevertheless, their odor is not a reliable means of detecting methylamines because the nasal membrane can be rapidly desensitized. The "MSA" Universal Tester with Detector Tube 92115 can be used to detect less than 10 ppm amines in air.

Methylamines are more basic than ammonia. They are very soluble in water, alcohols, ethers, and most polar solvents. At low temperatures, aqueous methylamines form crystalline hydrates: $CH_3NH_2 \cdot 3H_2O$, $(CH_3)_2NH \cdot 7H_2O$, and $(CH_3)_3N \cdot 10H_2O$. In the presence of ammonia, methylamines react explosively with mercury; so instruments containing mercury should not be used on methylamines. The amines also are flammable in air. Their flammability ranges as well as some of their important physical properties are listed in Tables 28.16 and 28.17.

Manufacture. Numerous routes are known to be feasible for the synthesis of methylamines. Among them, at least three routes are reported to exhibit a significant yield of

TABLE 28.16 Physical Properties of Anhydrous Methylamines

	MMA	*DMA*	*TMA*
Chemical formula	CH_3NH_2	$(CH_3)_2NH$	$(CH_3)_3N$
Molecular weight	31.06	45.08	59.11
Freezing point, °C	−93.5	−92.02	−117.3
Boiling point, °C	−6.3	6.9	2.9
Vapor density (1 atm, 25°C) g/liter	1.3	1.9	2.5
Specific gravity, air = 1	1.1	1.6	2.0
Liquid density (25°C), g/ml	0.6562	0.6496	0.6270
lb/gal	5.48	5.42	5.23
Liquid viscosity (25°C), cP	—	0.190	0.175
Critical temperature, °C	156.9	164.5	160.1
Critical pressure, atm	73.6	52.4	40.2
Enthalpy of formation, ΔH_f°(25°C), kcal/mol			
Gas	−5.49	−4.41	−5.81
Liquid	−11.3	−10.5	−11.0
Solution (1000 moles H_2O)	−16.78	−17.3	−18.6
Vapor pressure, psia			
25°C	50	30	32
100°C	400	230	205
Heat capacity, C_p(25°C), cal/°C · mol			
Gas	12.7	16.9	21.9
Liquid	—	32.9	32.31
Heat of fusion (m.p.), cal/g	47.20	31.50	26.46
Heat of vaporization (b.p.), cal/g	198.6	140.4	92.7
Autoignition temperature, °C	430	400	190
Flammable limits, vol. %			
Lower	4.9	2.8	2.0
Upper	20.7	14.4	11.6

Source: Du Pont Methylamines Bulletin.

TABLE 28.17 Physical Properties of Aqueous Methylamines

Property	*40% MMA*	*40% DMA*	*60% DMA*	*25% TMA*
Boiling point, °C	48	54	36	43
Freezing point, °C	−38	−37	−74.5	6
Liquid density (25°C)				
g/ml	0.897	0.892	0.829	0.930
lb/gal	7.49	7.44	6.92	7.76
Vapor pressure (25°C)				
psia	5.8	4.2	9.7	6.6
mm Hg	300	215	500	340
kPa	40	29	67	45
Flash point, closed cup, °C	−12	−18	−52	6

methylamines, worthy of industrial attention:

- Alkylation of ammonia by methanol or methyl ether.
- Reaction of carbon monoxide, ammonia, and hydrogen.
- Hydrogenation of hydrogen cyanide.

As of the early 1990s, practically all the commercial methylamine processes are based on the technology of methanol alkylation of ammonia.

Alkylation of Ammonia by Methanol. The alkylation reactions of ammonia originally were discovered by Sabatier and Maihle in the early 1900s for the synthesis of higher amines. The chemistry had been adopted for methylamines in the 1930s when mono-, di-, and trimethylamines (MMA, DMA, and TMA) were coproduced in the following vapor phase reactions in the presence of a dehydrating catalyst:[78]

$$\text{(I)}\quad NH_3 + CH_3OH \xrightarrow{\text{Cat.}} CH_3NH_2 + H_2O + 5.358\ \text{kcal/mol}$$

$$\text{(II)}\quad CH_3NH_2 + CH_3OH \xrightarrow{\text{Cat.}} (CH_3)_2NH + H_2O + 9.598\ \text{kcal/mol}$$

$$\text{(III)}\quad (CH_3)_2NH + CH_3OH \xrightarrow{\text{Cat.}} (CH_3)_3N + H_2O + 14.098\ \text{kcal/mol}$$

$$\text{(IV)}\quad NH_3 + (CH_3)_3N \xrightarrow{\text{Cat.}} CH_3NH_2 + (CH_3)_2NH - 8.74\ \text{kcal/mol}$$

$$\text{(V)}\quad NH_3 + (CH_3)_2NH \xrightarrow{\text{Cat.}} 2CH_3NH_2 - 4.24\ \text{kcal/mol}$$

$$\text{(VI)}\quad CH_3NH_2 + (CH_3)_3N \xrightarrow{\text{Cat.}} 2(CH_3)_2N - 4.50\ \text{kcal/mol}$$

The reaction mechanism and kinetics have been studied and described.[79] Although it was not possible to develop theoretical kinetic equations based on the Langmuir–Hinshelwood theory, the empirical kinetic expressions deduced from experiments using aluminum silicate catalyst are in the form of:

$$\frac{M}{A} = C_1 \exp[C_2 T + C_3 X]$$

for each of the amines where

M = g catalyst
A = moles MeOH fed per hour
T = reaction temperature
X = MeOH conversion
C_1, C_2, C_3 = coefficients corresponding to the ratio of MeOH to ammonia feed

Reactions (I) to (III), known as the alkylation reactions, are exothermic and highly irreversible except for reaction (III).[80] Reactions (IV) to (VI), known as disproportionation

reactions, are reversible and are endothermic if they proceed from left to right. The alkylation reactions dictate the rate of consumption of methanol and are somewhat faster than the disproportionation rates, which govern the selectivity of the three amines. However, the relative rate of alkylation and disproportionation depends on the type of catalyst and may be changed if a new catalyst is developed.

Thorium oxide was the original catalyst, but it was soon replaced by alumina and alumina silica catalyst[81] in the early 1930s because of their catalytic activity and selectivity. The improvement in the catalytic activity reduced the required reaction residence time and thus decreased the distribution of the three amines in the reactor effluent gases. Numerous patents have been issued, based on different recipes for preparing the catalyst for improvements in selectivity; they include U.S. Patent 3,387,032, E. German Patent 108.275, and Japan Patent 1623 (1952).[82-84]

Conventional catalysts consist of about 86 percent silica and 14 percent alumina and yield predominantly TMA, which is in the least commercial demand. The undesired TMA is separated and recycled to the reactor for disproportionation into other amines. U.S. patent 3,387,032 claimed that by adding 0.05 to 0.95 percent silver phosphate, rhenium heptasulfide, molysulfide, or cobalt sulfide to the conventional catalyst, with the proper steam treatment, the selectivity of the catalyst is shifted in favor of mono- and di- at the expense of tri-. The relative increases of MMA and DMA in the reactor effluent are 67 percent and 98 percent, respectively. The corresponding decrease in TMA is 63 percent.

Japan patent 1623 (1952) also claimed improvement in selectivity in favor of MMA and DMA with a MgO/SiO_2 catalyst (1:1 ratio). A more recent development in catalytic technology is covered in West German patent DT-2916-060 for the use of zeolite to produce predominantly MMA and DMA.[85] Although the cited patents, as well as many others, have claimed improvement in the catalysis for the synthesis of methylamines, no publication is available to show how many of these inventions actually are used industrially or to what extent.

Serban et al.[86-87] used the nitrogen/carbon (N/C)* mole ratio to correlate the effects of feed composition on selectivity. As expected, the MMA content in the reactor effluent increases as the N/C is increased, as the ammonia disproportionates TMA into lower amines. DMA, however, exhibits a maximum at an N/C of approximately 1, beyond which MMA begins to dominate. TMA, of course, decreases as the N/C increases.

Although the increase of the N/C can induce a more favorable selectivity, the benefit is counterbalanced by the recovery cost of unconverted ammonia. In addition to its dependency on the N/C ratio, the TMA content is reported to be reduced as water is fed to the reactor (i.e., the reversibility of Reaction III). The reaction temperature has a twofold effect on selectivity: increasing the temperature will increase the reaction rate of the disproportionation reactions; furthermore, a high temperature favors the equilibrium distribution of MMA and DMA in reactions (IV) to (VI). The published temperature dependency of the equilibrium constants is:

$$\frac{[\mathrm{MMA}][\mathrm{DMA}]}{[\mathrm{NH_3}][\mathrm{TMA}]} = \exp\frac{1725.1}{T} + 1.673 = K(\mathrm{IV})$$

$$\frac{[\mathrm{MMA}]^2}{[\mathrm{NH_3}][\mathrm{DMA}]} = \exp\frac{-1028.4}{T} + 0.877 = K(\mathrm{V})$$

$$\frac{[\mathrm{DMA}]^2}{[\mathrm{MMA}][\mathrm{TMA}]} = \frac{K(\mathrm{IV})}{K(\mathrm{V})}$$

where T is in °K.

The increase in reaction residence time also reduces the TMA content in the reactor effluent. Because the disproportionation reactions are the rate-limiting steps, an increase in residence time will favor the disproportion of TMA into lower amines.

In addition to the six primary reactions discussed above, formaldehyde, *N,N,N′,N′*-tetramethyldiaminomethane, and other trace impurities are known to be generated by side

*N/C was calculated as the moles of nitrogen in the feed divided by the moles of carbon.

reactions. Although the formation of by-products does not constitute a major yield loss, they present operating problems in the separation of amines. U.S. patent 3,720,715[88] showed that the injection of hydrogen into the reaction will suppress the side reactions.

A typical methylamines process is divided into three sections: synthesis, refining, and waste recovery (see Fig. 28.23).

In the synthesis step, the pressure is operated at 70 to 500 psig. Methanol, ammonia, and the recycled amines, which are in excess of market demand, are vaporized in a vaporizer. The mixture then is superheated to about 300 to 400°C before entering the reactor. The feed temperature, which is controlled to keep the MeOH conversion close to completion, is increased as the throughput in the reactor increases and as the catalyst deactivates during its life. The reactor typically is an adiabatic fixed-bed catalytic reactor where the residence time is 10 to 40 seconds.[82] The heat of reaction is recovered by preheating the feed mixture in a heat exchanger.

The temperature rise in the reactor depends on MeOH conversion and selectivity. The reactor exit temperature is typically at 400 to 500°C. The reactor effluent gases are condensed and stored for refining. Small amounts of noncondensable CO, H_2, and CH_4 are vented from the condenser. The yield of the reactions is over 90 percent for both ammonia and methanol. Although the MeOH conversion is controlled at close to 100 percent, the ammonia conversion is dependent on the N/C ratio. The ammonia conversion is typically at 10 to 40 percent. The crude product in a conventional process consists of:

Product	*Wt. %*
NH_3	20-50
MMA	7-12
DMA	10-22
TMA	10-30
MeOH	0- 3
H_2O	10-25

The technological challenge for improvement of the synthesis step is to reduce the amines recycle per pound of desired product.

Fig. 28.23. Leonard Process Co. process for methylamines. (Hydrocarbon Processing, Petrochemical Handbook, *p. 114, Nov. 1983. Copyright 1983 by Gulf Publishing Company and used by permission of the copyright owner.*)

The incentives are both a lower capital investment and lower energy consumption.

In the refining step, the amines are separated from the unreacted raw materials as well as from the reaction water by a series of distillations. Because the amines have a small differential in relative volatility and have an affinity for water, the fractions of the amines are very energy-consuming. Reportedly, it takes 10 to 30 lb of steam to produce 1 lb of amines. The energy efficiency of the process is dependent on the design of the equipment—which is a trade-off between the operating cost and the capital investment. Most of the distillation towers operate at pressures between 25 and 300 psig.

There are a least three reasons for distilling under pressure. First, high pressure increases the condensing temperature of the volatile amines and ammonia in the distillation overhead and thus eliminates the need for refrigerated condensers. Second, pressure increases the throughput capacity of a given tower diameter. Third, some of the ammonia–amine azeotropes disappear at elevated pressure. Azeotropic mixtures of ammonia–TMA, TMA–MMA, and TMA–DMA were reported in the literature.[89] Extractive distillation is also a viable way to overcome the difficulties of separating the amines. The extractive solvents are selected for their different solubilities of the three amines.

In 1983 the Leonard Process Company was licensing a continuous process for producing mono-, di-, and trimethylamines from technical-grade anhydrous ammonia and methanol. It was reported that during 1983 plants in 16 countries used the process to manufacture 220,000 tons of product.[90] The process (Fig. 28.23) is generally described as follows. The crude product from the converter goes through four distillation columns. The first separates excess ammonia and part of the TMA–ammonia azeotrope, which is recycled. In the TMA column water is added for extractive distillation, the TMA goes to storage, and the bottoms go to the third column, where pure MMA is recovered in the overhead. Then the bottoms go to the fourth column, where pure DMA is recovered overhead. It is claimed that all products are 99.5 percent pure, and that the yield on ammonia and methanol is over 95 percent.

UCB S.A. in Belgium has developed a continuous process for methylamines from ammonia and methanol, which is said to be pollution-free and for which the overall yield is 96.5 percent.[90]

In any refining scheme, the wastewater from the distillation tower is fed to a waste stripper where trace organics are removed prior to the sewer disposal. The organics are recycled to the process. In addition, the process and product storage vents are collected in a recovery system where amines and ammonia are recovered for recycle. Noncondensables are vented.

Other Routes. U.S. patent 3,444,203[91] describes the synthesis reactions of methylamines via carbon monoxide, ammonia, and hydrogen. The $CO/NH_3/H_2$ route has the advantage of backward integration to cheaper raw materials and also eliminates the expenditure for manufacturing methanol. Although the yields are reported to be in the range of 90+ percent, the process calls for high pressure equipment (e.g., 5000 psig) that results in a high capital investment that cannot compete with that of cheap methanol.

Reduction of hydrogen cyanide produces predominantly MMA. The chemistry is described in DBP 1,019,313[92] and others. This technology is not economically attractive except for backward integration to more elementary raw materials using HCN as an intermediate.

The reduction of formamide to methylamines also was published, in French Patent 762194.[93]

Economics and Marketing.[94,128] Methylamines are diversified in their uses. The patterns of use have changed only slightly in recent years, and the growth rate since the late 1970s has been only about 2 percent/year. About half of the monomethylamine had been used as a sensitizer in water gel explosives, but this use has declined substantially with the advent of less expensive explosives. This loss of demand, however, has been more than compensated for by the demand for its use in *N*-methylpyrrolidone (solvent). Dimethyl-

amine and trimethylamine growth has been near zero, and their use profile has not changed. Choline chloride is used as an animal feed supplement and in gastrointestinal medicines.[94]

The end uses for methylamines at the end of the 1980s are shown below:[77,95]

End Uses, Millions of lb/year

MMA		*DMA*		*TME*	
Explosives	18	Solvent		Choline	
N-Methylpyrrolidone	16	Dimethylformamide	30	Chloride	28
Pesticides	13	Dimethyl acetamide	5	Other	7
Alkylalkanoamines	7	Water treatment	20		35
Surfactants	3	Fatty tertiary amines	13		
Pharmaceuticals	1	Pesticides	9		
Other	2	Dimethylaminoethanol	9		
	60	Rubber processing chemicals	2		
		Other	7		
			95		

Methylamines are marketed as: the anhydrous grade; 25 percent and 40 to 60 percent aqueous MMA solutions; 25 percent and 40 percent aqueous DMA solutions; and 25 percent and 40 to 60 percent aqueous TMA solutions. The price for methylamines rose from 27 cents/lb in the early 1960s to 37 cents/lb (anhydrous grade) in 1988 and to 50 cents/lb in 1991. The major U.S. producers are listed below:[77,95]

Producer	*Capacity in million pounds per year in 1988*
Du Pont	180
Air Products	150
Rhone-Poulenc	22
Total	352

Ethylenediamine

Ethylenediamine (EDA) is the most important ethylene amine. In most conventional processes, EDA is coproduced with diethylenetriamine (DETA), triethylenetetramine (TETA), tetraethylene pentamine (TEPA), and pentaethylenehexamine (PEHA). Some of their physical properties are listed below:

	EDA	*DETA*	*TETA*	*TEPA*	*PEHA*
Molecular wt.	60.11	103.17	146.24	189.31	232.36
Boiling point, °C	11	−39	−35	−30	−26
Melting point, °C	117	207	277	decomp. @340°C	
Density, lb/gal	7.67	7.96	8.19	8.33	8.39
Flash pt. (O.C.), °C	38	102	143	185	185

Ethylenediamine is a slightly viscous liquid with a strong odor of ammonia. It is soluble in water, ethanol, acetone, ether, and benzene. The threshold odor limit for EDA is 10 ppm.

Commercial ethylenediamine is produced by two different routes:

- From ethylene dichloride (EDC) and ammonia.
- From ethylene oxide (EO) and ammonia via monoethanolamine (MEA).

In the United States, about 60 percent of the ethylene amines are produced by the EDC route. The EDC and MEA routes both are employed by Union Carbide in its facilities at Taft, Louisiana. Both of these routes produce other higher amines (e.g., DETA, TETA, TEPA, and PEHA) as well; but the MEA route produces predominantly ethylenediamine. In Europe, approximately 40 percent of the ethylenediamine is produced by the MEA route, and the balance by the EDC route.

Ethylene Dichloride and Ammonia. This process reacts EDC with aqueous ammonia to produce EDA and higher homologues. The

primary reaction is:

$$ClCH_2CH_2Cl + 2NH_3 \rightarrow NH_2CH_2CH_2NH_2 + NH_2CH_2CH_2NHCH_2CH_2NH_2 + \text{Highers}$$

Other by-products are triethylenetetramine, tetraethylenepentamine, pentaethylenehexamine, and polyethylene polyamines. A low reaction temperature and low pressure are more selective for EDA, but conversion is low. High temperature and high pressure, on the other hand, increase conversions, but with less selectivity for EDA. The reaction system is noncatalytic. In a typical EDA process, ethylene dichloride reacts with excess aqueous ammonia at about 100 to 200°C and at 10 to 20 atmospheres. The crude product is an aqueous solution of ammonia, ammonium chloride, ethylenediamine hydrochloride, and higher amine hydrochlorides. Caustic is added to neutralize the crude product and free the ethylenediamine, the higher amines, and the ammonia from the amine hydrochloride and ammonium chloride. The chloride salts must be disposed of. The crude product is refined by fractionation, in which the unreacted ammonia is recycled and the amines are purified. The EDA yield is estimated to be between 45 and 70 percent.[96,97]

Monoethanolamine and Ammonia. This reaction takes place at high pressure over an amination catalyst such as nickel or cobalt–nickel supported on alumina, titania, or a similar metal oxide. The primary reaction is:

$$\underset{\text{Monoethanol amine}}{NH_2CH_2CH_2OH} + NH_3 \rightarrow \underset{\text{Ethylenediamine}}{NH_2CH_2CH_2NH_2} + H_2O$$

Coproducts DETA and other highers and cyclic amines are produced simultaneously. The ammonia-to-MEA ratio ranges from 2:1 to 25:1 (mole basis) with hydrogen present to the extent of up to 25 percent by volume to maintain a reducing atmosphere over the catalyst. The liquid phase reaction temperature is 220 to 270°C at 800 to 4500 psig. The crude product is refined by fractionations. The MEA usually is produced from ethylene oxide and ammonia.

Economics and Marketing. The fastest-growing end use of ethylenediamine is in the manufacture of the chelating agent EDTA (ethylenediaminetetraacetate), which is a product of the reaction of EDA with chloroacetic acid or sodium cyanide/formaldehyde. Its average growth rate in the late 1980s was about 6 percent.

Sharing approximately the same proportion of EDA markets is the manufacture of polyamide resins, for hot melt adhesives and other uses, and ethylene bis(stearamide), used as a mold release agent. A small proportion of EDA is used to manufacture dimethylolethyleneurea resins for permanent-press and easy-care fabric finishes in the textile industry. The export market has occupied a substantial share of EDA production.

The distribution of EDA among end uses is shown below. The average growth rate through the late 1980s was about 5 percent.

Annual U.S. EDA Use, millions of pounds

Polyamide resins	10
Ethylene bis(stearamide)	9
Chelating agent	22
Miscellaneous surfactants	4
Export	40
Gasoline and lube oil additives	6
Other	10
	101

The pricing history of EDA is shown in Fig. 28.24. The major U.S. producers are:

*1989 Capacity**

Union Carbide	180
Dow	100
Air Products	20
Texaco	15
Total	315

*Millions of pounds per year of ethyleneamines.

A process licenser for Methylamines is UCB.[47]

165, 166, 167, 26

Fig. 28.24. Pricing history of ethylenediamine. (Chem. Econ. Handbook *648,5032 and 651,5032*; Chem. Mktg. Rept. *1991*.)

HEXAMINE

Hexamine, $(CH_2)_6N_4$, is also known as hexamethylenetetramine, aminoform, crystamine, methenamine, or formin. It was first prepared in 1859 by Butlerov of Russia. Hexamine has a three-dimensional molecular structure with the carbon atoms occupying the peaks of an octahedron and the nitrogen atoms occupying the peaks of a tetrahedron that circumscribes the octahedron as shown in Fig. 28.25.

Hexamine is a colorless solid at temperatures below 285°C, and has a crystalline structure of rhombic dodecahedrons. It does not melt, but sublimes at 285 to 295°C. Odorless with a sweet metallic taste, hexamine is known to cause skin rashes on some people who are exposed to its vapor or solid form. It is flammable and burns with a bluish-yellow flame.

Thermal decomposition of hexamine produces ammonia, hydrogen cyanide, methane, nitrogen, hydrogen, and residue oils. It also hydrolyzes into ammonia and formaldehyde in acidic solutions. Hexamine reacts with hydrogen peroxide in acid solution to form hexamethylenetriperoxide diamine (HMTP), which is an explosive. When nitrated, hexamine is converted into cyclotrimethylene-trinitramine, also known as hexahydro-1,3,5-trinitro-s-triazine, cyclonite, hexagen, or RDX. It was an ingredient of the so-called blockbuster bombs, which were listed among the most powerful explosive devices in World War II.

An aqueous solution of hexamine exhibits the phenomenon of inverse solubility. At 25°C

Fig. 28.25. Three-dimensional schematic of hexamine.

water can dissolve up to 0.867 g hexamine/cc, whereas at 50°C the solutility is decreased to 0.807 g/cc. Hexamine in aqueous solution demonstrates the properties of a weak monobase with pH ranging from 8 to 9. The hydrate, $(CH_2)_6N_4 \cdot 6H_2O$, of hexamine can be crystallized from aqueous solution at temperatures below 14°C. Some of its physical properties are listed below:

Molecular weight	140.19
Sublimation temperature	285–295°C
Density @ −5°C	1.331 g/cc
Specific heat	36.5 cal/°C
Heat of formation @25°C	28.8 kcal/mol
Heat of Combustion @25°C	1003 kcal/mol
Free energy	102.7 kcal/mol
Entropy	37.048 cal/°C
Solubility @ room temperature,	
Chloroform	0.134 g/cc
Methanol	0.073 g/cc
Ethanol	0.029 g/cc
CCl_4, acetone, benzene, ether	<1%

Manufacture

Hexamine is manufactured by the liquid phase reaction of ammonia and formaldehyde:

$$4NH_3 + 6HCOH \rightarrow$$

$$(CH_2)_6N_4 + 6H_2O + 55 \text{ kcal/mol hexamine}$$

The reaction is controlled at pH 7–8 because formic acid and carbon dioxide are formed under acidic conditions.

Early commercial production of hexamine used a batch reaction of an aqueous solution of formaldehyde and ammonia at ~30 percent and ~20 percent concentration, respectively. The reaction temperature was controlled at ~20°C via a reactor cooler to remove the heat of reaction. With the low reaction temperature and the relatively high exotherm, temperature control was difficult as well as expensive because of the large cooling surface required. The hexamine concentration in the crude product was about 15 percent by weight. The aqueous hexamine solution was concentrated and crystallized under a 40 mmHg vacuum. Approximately 5500 Btu/lb of hexamine produced were required to evaporate 5.7 lb of water/lb of hexamine in the crude product solution. Ammonia was added during the evaporation step to avoid hydrolysis of hexamine. The batch cycle ranged from 8 hours to over a day, depending on the design criteria.

Figure 28.26 is a schematic representation of a continuous hexamine process. Here, anhydrous ammonia is used to reduce the amount of water entering the reactor and thus reduce the heat requirement in the evaporation step. Industrial grade formaldehyde usually is available in 37 to 50 percent aqueous solution. The reacting solution is controlled at 30 to

Fig. 28.26. Manufacture of hexamethylenetetramine. (*Adapted from* Petrol. Ref., ***37**, No. 9, 351–353, 1958.*)

50°C by being circulated through an external water-cooled heat exchanger. The residence time in the reactor is in the range of 15 to 30 minutes. Some processes use gaseous ammonia instead of liquid ammonia as a raw material. In that case, the reactor cooler must be capable of accommodating the 60 percent additional heat generated from the heat of solution of the gaseous ammonia. If gaseous ammonia is used, the heat load in the reactor cooler is approximately 88 kcal/mol of hexamine versus 55 kcal/mol if liquid ammonia is used. To avoid local hot spot and hammering effects of mixing gaseous ammonia and formaldehyde, a stoichiometric amount of the reactants is added to a large circulating stream of hexamine solution.

Although the use of liquid ammonia has the advantage of causing a smaller heat load in the reactor cooler, it has the disadvantage of product contamination caused by impurities in the liquid ammonia. But whether the ammonia is introduced to the reactor in liquid or gaseous form, the use of anhydrous ammonia reduces the water content in the crude product solution by 35 percent and increases the hexamine concentration to 21 percent.

The reactor effluent is fed to a vacuum crystallizer where the product is concentrated and crystallized. Approximately 3700 Btu/lb of hexamine produced is required to evaporate 3.8 lb of water/lb of hexamine in the crude product solution. The slurry is centrifuged, and the crystals are washed and discharged to a dryer. The mother liquor is recycled to the reactor system, but a small bleed is necessary in order to avoid the buildup of impurities in the system. A small amount of ammonia is continuously added during the evaporation step to reduce the decomposition of the hexamine. If the formaldehyde feed contains methanol, it will be removed from the evaporator system along with the ammonia and water. The temperature in the final drying step is held to below 50°C to avoid yield losses due to decomposition. The overall yield is in the range of 95 to 96 percent based on formaldehyde; that is, 3.64 lb of formaldehyde are required to manufacture 1.0 lb of hexamine. The reactor usually is made of stainless steel or aluminum, and the reactor cooler and the evaporator are made of stainless steel.

The Meissner process (Germany) for the manufacture of hexamine, still widely used today, was introduced in the early 1950s. This process [98,99] uses the heat of reaction and the heat generated from the mixing of ammonia and formaldehyde in the reactor to vaporize the water from the reaction. The reactants are introduced to the reactor as gaseous ammonia and aqueous formaldehyde. The reactor temperature is controlled at 50 to 70°C by regulating the total pressure at which the reaction mixture is allowed to boil or by varying the partial pressure of inert gases. The amount of water vaporized corresponds to the heat generated in the reactor. Reportedly, the hexamine solution can be concentrated to 25 to 30 percent in a reactor that has the configuration of a continuous vacuum distillation column. If additional heat is added to the reactor, hexamine can be crystallized in the reactor, thereby eliminating the need for a crystallizer. To crystallize the hexamine, a total of 3700 Btu/lb of product is required for removal of the water of reaction and the water in the formaldehyde solution (3.8 lb H_2O lb hexamine). However, the heat of reaction and of the mixing of ammonia and formaldehyde supplies approximately 1100 Btu/lb of hexamine (i.e., 30% of the heat required); so a net heat input of 2600 Btu to the reactor/lb of hexamine is required if the product is to be crystallized in the reactor. The overhead vapor from the reactor is condensed, and the methanol and other volatile impurities are removed from the system as off-gas. Reportedly, the reaction goes to 98 percent completion based on both reactants.

In 1970, Tenneco, Inc. patented a hexamine process[100] in which the hexamine is separated from the reaction mixture via a spray dryer instead of an evaporator. In this process, gaseous ammonia is reacted with a 30 to 50 percent aqueous formaldehyde solution in approximately a 0.67 NH_3/1.0 formaldehyde (mole) ratio. The formaldehyde solution is prepared from deionized water and thus is

low in cation content. The reaction mixture is maintained at 50 to 90°C and at a pH of 7.5 to 8. The reaction residence time ranges from 2 to 5 minutes. The aqueous reaction solution (25–35 wt. % hexamine) is fed continuously into a spray dryer, in which it is atomized and is contacted with a stream of air or inert gas preheated to 200 to 400°C. The hexamine in the spray is solidified and dehydrated to 0.1 to 0.2 wt. % of water. The exit temperature of the inert gas is maintained below the decomposition temperature of the product (at 100–120°C). The typical residence time of the drying step is 0.5 to 5 seconds. The hexamine particles then are separated from the inert gas and subjected to 1000 to 2500 psi pressure in a pair of cast steel rolls where sheets of densified hexamine are formed. In the final step the sheets are subdivided into dust-free granules ranging in size from 100 to 300 microns.

A Czechoslovakian patent issued in 1988[101] claims significant improvement to established processes by using as the wash medium the liquefied condensate obtained by evaporating aqueous hexamine solutions in amounts of 0.01 to 1 kg/kg of crystalline hexamine.

Economics and Marketing

Military use of hexamine in the manufacture of cyclonite accounted for most of its rapid growth in the United States in the second half of the 1960s. When nitrated into cyclotrimethylene-trinitramine, hexamine is converted into an ingredient for the production of the powerful cyclonite explosives. The military consumption of hexamine has been severely curtailed since the mid-1970s, except during the Persian Gulf warfare involving the United States in 1991. The principal uses of hexamine in a peacetime economy are:

- As a curing agent in thermosetting resin production. It serves as a methylating agent in the curing of phenol-formaldehyde resins.
- As a chelating agent. The production of nitrilotriacetic acid and its salt is the second largest peacetime consumption of hexamine and appears to be growing.
- As an accelerator. In the rubber industry it is used to prevent vulcanized rubber from blocking.

Other uses for hexamine are in pharmaceuticals, as an inhibitor of corrosion caused by strong mineral acids, for shrink-proofing agents in the textile industry, and as an agent to enhance color fastness and elasticity to cellulose fibers. It also has potential for end use as a fungicide in the citrus fruit industry. Hexamine produces methylene groups without forming water, as in the case with paraformaldehyde.

The list price for hexamine in 1991 was 70 cents/lb granular in bags.[46] Licensers for the process are Karl Fischer and Josef Meissner.[47]

The major U.S. manufacturers are:

*Hexamine Manufacturer**	*1990 Capacity in million PPY*
Borden Inc.	30
Plastics Eng. Company	8
Wright Chemical Corps	40
Total	78

*W. R. Grace & Co. and Monsanto Co. produce HMTA as an unisolated intermediate for captive use in the production of nitriloacetic acid.[103]

HYDRAZINE

Hydrazine, $H_2N—NH_2$, is the simplest diamine. Free hydrazine never has been detected in nature on earth although many hydrazine derivatives have been found. Anhydrous hydrazine was prepared first by Lobry de Bruyn in 1894, In 1938, Fairmont Chemicals began the first production of hydrazine in the United States; the quantity produced was small, and the product was mainly for captive use in boiler feedwater treatment. In World War II, hydrazine became important to Germany as a component of missile fuel. In 1954, Olin Chemicals built the first U.S. commercial N_2H_4 plant, of about 4 million lb/year capacity. Hydrazine mainly is consumed as a high energy rocket fuel, an oxygen

scavenger for boiler feedwater treatment, and in blowing agents for polymer industries, agricultural chemicals, and pharmaceuticals; but during the past decade its rocket fuel use has declined. It also has a promising future as a fuel for fuel cells. Because of its dual active nitrogen sites, hydrazine and its derivatives are versatile tools for building heterocyclic compounds.

Hydrazine is a clear, hygroscopic fuming liquid with the odor of ammonia. It is a mild base and is highly polar, being miscible in polar solvents such as water, ammonia, amines, and alcohols. Its high dielectric constant accounts for its unusually high boiling point (113.5°C) and makes it a good solvent for ionic reactions. In the liquid phase, hydrazine is known to be associated although it is not associated in the vapor phase. It is believed to be in gauche form and exhibits a dipole moment, forming hydrogen bonds with water. It forms an azeotropic mixture with water at 68 percent hydrazine with a boiling point of 120.5°C. Hydrazine hydrate (H_2NNH_3OH) is obtained by fractional distillation from water. The monohydrate has a boiling point of 119°C, a melting point of −40°C, and a density of 1.03 g/ml at 21°C. The physical properties of hydrazine are listed in Table 28.18.

TABLE 28.18 Physical Properties of Hydrazine

Molecular weight	
Anhydrous	32
Hydrate	50
Boiling point (°C)	113.5
Melting point (°C)	1.4
Specific gravity (g/ml)	
@0°C	1.025
@15°C	1.014
@25°C	1.004
@50°C	0.982
Critical temperature (°C)	380
Critical pressure (atm)	145
Vapor pressure (mm Hg)	
@25°C	14
@31°C	20
@36°C	100
Viscosity (cp)	
@5°C	1.2
@25°C	0.9
Heat of vaporization (kcal/mol)	9.6
Heat of solution (kcal/mol) @25°C	−3.9
Heat capacity (cal/mol)@25°C	23.6
Heat of combustion (kcal/mol)	−146.6
Heat of formulation (kcal/mole)	
Liquid	12
Gas	23
Flash point (°C)	52
Explosive limits in air by vol. %	4.7–100

Hydrazine is a strong reducing agent. When catalyzed by metal, it decomposes into nitrogen, ammonia, and hydrogen:

$$\underset{\text{Hydrazine}}{3NH_2NH_2} \rightarrow 4NH_3 + N_2$$

$$NH_2NH_2 \rightarrow 2H_2 + N_2$$

The four hydrogen atoms in hydrazine molecules can be substituted for to form monosubstituted $RNHNH_2$, unsymmetrical disubstituted R_2NNH_2, symmetrical disubstituted RNHNHR, trisubstituted R_2NNHR, and tetrasubstituted R_2NNR_2. Alkylhydrazines, such as monoalkylhydrazine and dialkylhydrazine, resemble hydrazine in their properties. They are weak bases, strong reducing agents, and water-soluble. In fact, monomethyl and unsymmetrical dimethyl hydrazines are listed with hydrazine among the highest-performing high energy fuels, whereas tri- and tetra-alkyl hydrazines are relatively weak reducing agents.

Hydrazine vapor is flammable in air in the range of 4.7 to 100 vol. %. Liquid hydrazine burns like gasoline when ignited freely in air although it burns fiercely and at an elevated temperature. An aqueous solution below 40 percent is not ignitable. Small spills of hydrazine are best handled by deluging them with large quantities of water. Hydrazine is not detonable by shock.[104]

Hydrazine can be fatally toxic via injection, ingestion, inhalation of vapors, or contact with skin. Exposure to hydrazine causes irritation of the eyes, nose, and throat, and leads to dizziness and nausea. The symptoms may appear gradually over several hours of exposure. Although hydrazine has a relatively low threshold odor limit (70–80 ppm), smelling is not a reliable detection for its presence

because the membranes of the nose may be desensitized rapidly. The threshold limit value adopted by the American Conference of Governmental Industrial Hygienists is 0.1 ppm, as it is a suspected carcinogen.

Manufacture

Hydrazine now is produced commercially by a reaction that differs greatly from the reaction used at its discovery when diazoacetic ester was hydrolyzed with alkali. Attempts to make hydrazine by direct synthesis from the elements have been unsuccessful on a practical basis. Three routes for the synthesis of hydrazine are of industrial significance. The Raschig (Olin) process and the Hofmann (urea) process are the conventional routes to the manufacture of hydrazine. The third method is a novel process developed by Produits Chimiques Ugine Kuhlmann of France. The Raschig process is presently the principal process for large-scale manufacture of hydrazine. Ammonia and sodium hypochlorite are the raw materials, with chloramine as an intermediate. The Hofmann process substitutes urea for ammonia as raw material and thus eliminates the need for ammonia recovery facilities. Thus, although the raw material cost in the urea process is somewhat higher than that of the Raschig process, the urea process has the advantage of lower investment cost. Both the Raschig and the Hofmann processes produce sodium chloride as a by-product and thus require investment in facilities, as well as energy, for the crystallization of NaCl.

The Ugine Kuhlmann process uses hydrogen peroxide instead of sodium hypochlorite as a raw material and eliminates the capital and energy costs for crystallizing the by-product NaCl produced in the conventional processes. The capital and the manufacturing costs are estimated to be reduced by 10 percent with the Ugine Kuhlmann process.

Raschig and Olin Processes. The Raschig process originally was discovered in 1907 and subsequently was modified into the Olin process.[105] The chemical reactions take place in the liquid phase and are divided into three steps:

$$(1)\ NaOH + Cl_2 \rightarrow \underset{\text{Sodium hypochlorite}}{NaOCl} + HCl$$

$$(2)\ NH_3 + NaOCl \rightarrow \underset{\text{Chloramine}}{NH_2Cl} + NaOH$$

$$(3)\ NH_2Cl + NH_3 + NaOH \rightarrow \underset{\text{Hydrazine}}{N_2H_4 \cdot H_2O} + NaCl$$

Hydrazine is obtained in the form of hydrated hydrazine with one mole of water added. Although a significant fraction of hydrazine is used as the hydrate, numerous applications such as rocket propulsion require anhydrous hydrazine. Above 64 percent N_2H_4, water is not easily removed. Chemical reaction followed by distillation or azeotropic distillation with another fluid is needed. Precautions must be taken to avoid explosive vapor decomposition.

The side reactions are:

$$(4)\ 2NH_3 + \underset{\text{Chloramine}}{3NH_2Cl} \xrightarrow{Cu^{++}} 3NH_4Cl + N_2$$

$$(5)\ 2NH_2Cl + N_2H_4 \rightarrow 2NH_4Cl + N_2$$

Protein-like materials such as albumin and glue are used as inhibitors, by forming complexes with the metal ions that otherwise would promote the side reactions.

In Reaction 1, the formation of sodium hypochlorite is spontaneous with *in situ* mixing of chlorine and aqueous caustic solution. In Reaction 2, the reaction rate for chloramine formation is rapid relative to the formation of hydrazine in Reaction 3. Because the decomposition rates are relatively insensitive to temperature, Reaction 3 is operated at 130°C to speed up the rate-determining step. Excess ammonia is used to minimize the hydrazine–chloramine decomposition.

Figure 28.27 is the process schematic of the Raschig process. Sodium hypochlorite is produced by continuous feeding of chlorine into aqueous sodium hydroxide (30%) in a cooler-reactor system.

The temperature is kept below 30°C in order to prevent the formation of sodium chlorate.

Fig. 28.27. Hydrazine manufacture. (*Adapted from Chilton, C. H.,* Chem. Eng., ***65****, No. 14, 123, July 14, 1968.*)

For the same reason, the sodium hydroxide concentration is kept below 1 g/liter. Excess caustic will overload the subsequent crystallization step. The inhibitor for side reactions such as glue is added to the solution until the mixture is viscous. In the chloramine reactor, dilute ammonia solution (5–15%) is added to the hypochlorite solution at a ratio of 3:1. Excess anhydrous ammonia then is added to the chloramine solution in the hydrazine reactor at a ratio of 25:1 at 130°C. Synthesis efficiency favors a dilute system although the increase in operating cost due to the low concentration may ultimately become inhibiting.

In the first step of refining, the unreacted ammonia is removed from the reactor effluent in the overhead of the ammonia stripper. The tails are fed into an evaporator where sodium chloride is concentrated and removed. The vapor from the evaporator is dehydrated in the three subsequent columns. In the first dehydrator, water is removed until the N_2H_4–H_2O mixture approaches the azeotropic concentration of about 65 percent hydrazine. In the second dehydrator, aniline is used as the extracting agent to break the azeotrope, and the water is removed in the overhead with the aniline, which is subsequently recovered in a decanter for recycle. The final column removes aniline from the hydrazine, producing 98 percent purity N_2H_4.

Vacuum distillation is used to reduce decomposition. However, the leakage of air into the process may result in decomposition of hydrazine into nitric oxide, ammonia, and water. In some cases, anhydrous hydrazine is obtained by dehydrating with 50 percent caustic.

The overall yield based on chlorine is 65 percent. The combined yield for Reactions 1 and 2 is about 95 percent, whereas the hydrazine step has a yield of 70 percent.

A 1987 Japanese patent publication disclosed a method of increasing the hydrazine yield by almost 10 percent by high frequency heating immediately after adding and mixing hypochlorite solution with ammonia water.[106] The improvement is due to the rapidity of the temperature elevation.

The Raschig process also can be extended to react amines with chloramine to produce monosubstituted or unsymmetrical disubstituted hydrazines; for example:

$$NH_2Cl \xrightarrow{NH_3} H_2N{=}NH_2$$
$$NH_2Cl \xrightarrow{RNH_2} RNHNH_2$$
$$NH_2Cl \xrightarrow{R_2NH} R_2NNH_2$$

Hoffmann (Urea) Process. The overall reaction proceeds in liquid phase as:

$$\underset{\text{urea}}{NH_2\overset{\overset{O}{\|}}{C}NH_2} + NaOCl + 2NaOH \rightarrow$$

$$N_2H_4 + NaCl + Na_2CO_3 + H_2O$$

Fig. 28.28. Hoffman process for hydrazine.

As in the Raschig process, aqueous caustic is chlorinated into sodium hypochlorite solution (Fig. 28.28). The urea solution is prepared by dissolving urea in water with the addition of steam to override the endothermic dissolution. The temperature is kept at about 5°C for 43 percent urea solution. Glue is added as an inhibitor to the side reactions at a ratio of 0.5 g/liter of solution; the urea and hypochlorite solutions are added to the hydrazine reactor at a ratio of 1:4. The reaction temperature is allowed to rise to 100°C. The crude product contains approximately 35 g N_2H_4/liter and can be refined in the same procedure as used for the Raschig process.

Ugine Kulmann Process. The technology for this process was developed by Produits Chemiques Ugine Kulmann of France, based on the ammonia, hydrogen peroxide, and ketone route to hydrazine. Methyl ethyl ketone is present in the feed to the reactor, and it participates in the reaction sequence; but in the final sequence it is reformed, and the net reaction is $2NH_3 + H_2O_2 \rightarrow N_2H_4 + 2H_2O$, with the ketone being recycled.

The schematic of the process is shown in Fig. 28.29. The manufacturing procedure is divided into five major steps: the synthesis of azine, the removal of low boilers, high boilers separation, hydrolysis of azine, and the dehydration of hydrazine. The synthesis of azine requires a residence time of 7 hours. Ammonia, hydrogen peroxide, and methyl ethyl ketone (MEK) are fed into the reactors along with methanol and acetonitrile at about 50°C. Small amounts of the disodium salt of ethylenediamine tetracetic acid are added to the reactants. In the low-boilers removal step, the unreacted ammonia, MeOH, MEK, and acetonitrile are separated and recycled to the reactors. The high-boilers solution contains water, MEK–azine, and acetamide. The acetamide is formed in the following reaction:

$$2NH_3 + H_2O_2 + \underset{\text{MEK}}{CH_3\overset{O}{\overset{\|}{C}}CH_2CH_3} + \underset{\text{Acetonitrile}}{CH_3CN} \longrightarrow \underset{\text{An azine}}{\begin{matrix} CH_3 & & CH_3 \\ \diagdown & & \diagup \\ & C{=}N{-}N{=}C & \\ \diagup & & \diagdown \\ CH_3CH_2 & & CH_2CH_3 \end{matrix}} + \underset{\text{Acetamide}}{CH_3\overset{O}{\overset{\|}{C}}NH_2} + 3H_2O$$

Fig. 28.29. Schematic of Ugine Kuhlman process for hydrazine.

After being separated from the solution as a high boiler, the acetamide is dehydrated into acetonitrile at a high temperature by using a catalyst of orthophosphoric acid–impregnated alumina. The acetonitrile is recycled to the reactors. In the hydrolysis step, the aqueous MEK–azine solution is partially decomposed into an aqueous solution of MEK and hydrazine.

MEK is separated and recycled to the synthesis step, and the hydrazine is fractionated from the unhydrolyzed aqueous azine solution, which is reworked through the hydrolysis step. The hydrazine is marketed as the hydrate or dehydrated with aniline.

The estimated yield is 42 mole percent ammonia and 73 mole percent hydrogen peroxide.

Other Technologies. There are several other processes, which are not used commercially:[107]

The Bergbau or Bayer Process is somewhat similar to the Ugine Kuhlmann technology where a ketazine intermediate is formed. The

reaction routes are:

This process has not been used on a commercial scale.

Direct synthesis of hydrazine is reported via passage of ammonia through an extended glow discharge but is not commercially feasible. The reaction is:

$$2NH_3 \rightleftarrows N_2H_4 + H_2$$

Highly turbulent flow and effective cooling are required. Allyl compounds are used to increase the yield.

Anhydrous hydrazine also is formed by the reaction of chlorine and ammonia:

$$Cl_2 + 4NH_3 \rightarrow N_2H_4 + 2NH_4Cl$$

The reaction sequence is postulated to be:

(1) $Cl_2 + 2NH_3 \rightarrow NH_2Cl + NH_4Cl$
(2) $NH_2Cl + NH_4Cl + 2NH_3 \rightarrow N_2H_4 + 2NH_4Cl$

High temperature and high pressure increase the reaction rate. A large ammonia to chlorine ratio (50:1 to 350:1) is required to obtain an acceptable yield.

A licenser for production of hydrazine is reported to be the Andhra Sugars Co.[47] The raw materials are caustic soda, chlorine, and urea.

Some uses require ultrapure hydrazine. The main contaminants are aniline and water. Hydrazine used as a propellant can generate hydrogen cyanide in the exhaust gases; so it is important to remove carbon compounds. The usual method is freeze crystallization, where the supernatant fluid is removed. Assays between 99.5 and 99.99 percent have been achieved.

Hydrazine Handling

The handling of hydrazine has been discussed extensively by Schmidt.[108] For structural applications in tanks, pipelines, valves, and so on, the metals most commonly used are alloys of titanium, aluminum, and stainless steels (300 series). Teflon®, neoprene, and nitrile rubber have been used.

Hydrazine is classified by the DOT as a corrosive liquid and is shipped in 40,000- and 50,000-lb quantities in tank trailers and tank cars as well as in 440-lb drums. It normally is stored under a blanket of nitrogen to avoid water and CO_2 contamination.

Protective equipment consists of air masks, gloves, boots, aprons, and splash shields. The gloves are made of butyl or neoprene. Chemicals used to decontaminate diluted hydrazine in water are calcium hypochlorite, sodium hypochlorite, and hydrogen peroxide. Strong hydrazine can be burned as a disposal method.

Economics and Marketing

The versatile end uses of hydrazine reflect its attractive properties as a strong reducing agent and a source of powerful energy. The consumption of hydrazine as a high energy fuel was the dominating market through the 1960s. In the 1970s, end uses in the manufacture of agricultural chemicals became the major customer for hydrazine, and they were the fastest-growing market in the 1980s. In spite of its versatile chemical properties, the relatively high price of hydrazine ($1.35/lb in drums) may be a hindrance to its growth. The end use pattern in the mid-1960s is compared

with the patterns in the late 1970s and mid-1980s below:

	Mid-1960s Basis = 15 million ppy consumed (pure hydrazine)	*Late-1970s Basis = 20 million ppy consumed (pure hydrazine)*	*Mid-1980s*
High energy fuels	73%	20%	5%
Agricultural chemicals	11%	35%	40%
Blowing agents	8%	25%	33%
Boiler water treatment	3%	10%	15%
Pharmaceuticals	3%		3%
Miscellaneous	2%	10%	4%

Hydrazine is marketed as anhydrous hydrazine, as the monohydrate (64% hydrazine), and as a 54.5 percent aqueous solution (85% hydrazine monohydrate). The current major U.S. producers are Mobay Chemicals and Olin Chemicals.

Agricultural Chemicals. The fastest-growing application of hydrazine has been in the manufacture of herbicides, pesticides, fungicides, and plant growth regulators. In 1980s, this application accounted for 40 percent of the total consumption of hydrazine. For example, as a plant-growth regulator, maleic hydrazine interferes with the metabolism of the growth tips of some root crops and grasses, and thus temporarily inhibits their growth. It is used to prevent the sprouting of potatoes and onions in storage, to control the growth rate of grass so as to reduce the frequency of mowing in public parks, and to combat the growth of suckers on tobacco plants. It also has many other usages.

The N—N linkage of some of the hydrazine derivatives adds to their activity as agricultural chemicals because of their self-degradation mechanism. Thiadiazole is an example of this unique property, where the biological activeness of the compound is complemented by its self-destruct mechanism.

Blowing Agents. Blowing agents such as 2,2′-azoisobutyronitrile, benzene sulfonyl hydrazide, and azodicarbonamide are used for the manufacture of foam plastics. These compounds decompose at elevated temperature (e.g., $>100°C$) to yield large quantities of nitrogen (e.g., >100 ml/g) together with some other compounds. Azodicarbonamide has the advantage of decomposing into nontoxic compounds in addition to the desired nitrogen. However, the market for blowing agents has matured.

High Energy Fuels. Hydrazine, monomethylhydrazine, and unsymmetrical di-methylhydrazine are classified as high-performing liquid rocket fuels because of their high specific density impulses and spontaneous combustion. The mono- and dimethylhydrazines have an advantage over hydrazine because of their lower freezing points ($-50°C$ vs. $2°C$). The rocket fuel markets constituted 20 percent of hydrazine consumption in the late-1970s, but fluctuation in U.S. defense budgets and funding of the space missile programs has from time to time varied the demand for hydrazine. Safety concerns over the storage of fuel have diminished its attractiveness.

Boiler Feedwater Treatment. The strong reducing characteristic of hydrazine makes hydrazine an oxygen scavenger for the treatment of boiler water to reduce corrosion, with the addition of 0.1 ppm.

$$NH_2NH_2 + O_2 \rightarrow 2H_2O + N_2$$

This application, which appears to have matured, is by far the most effective treatment for boiler feedwater because the treatment does not produce solid residues and avoids sulfide corrosion. Hydrazine also helps to form an adherent layer of magnetite, Fe_3O_4,

over the iron surfaces of the boilers and thereby reduces corrosion. However, in spite of its efficiency as an oxygen scavenger, the high price of hydrazine has limited its growth in the boiler feedwater treatment market. Also concern about possible carcinogenicity of hydrazine has curtailed its use in steam-heated food and drug processing facilities.

Miscellaneous. In addition to its versatile chemical nature for the manufacture of pharmaceuticals, hydrazine has a high potential for growth as fuel for fuel cells because of its reactivity:

Anode: $NH_2NH_2 + 4OH^- \rightarrow N_2 + 4H_2O + 4eE^\circ = 1.16V$

Cathode: $O_2 + 2H_2O + 4e \rightarrow 4OH^-$ $E^\circ = 0.4V$

Overall: $NH_2NH_2 + O_2 \rightarrow N_2 + 2H_2O$ $E^\circ = 1.56V$

Hydrazine is a useful reductant for electrodeless plating baths for copper, nickel, cobalt, and silver.

HYDROGEN CYANIDE

Hydrogen cyanide (HCN) is also known as hydrocyanic acid, prussic acid, or formonitrile. It is a very important building block used in producing adiponitrile, acetone cyanohydrin, sodium cyanide, cyanuric chloride, chelating agents, methionine, and other intermediates. These intermediates, in turn, serve as raw materials for the manufacture of synthetic fibers, plastics, herbicides, dyestuffs, pharmaceuticals, explosives, surfactants, and metal treatment agents. A detailed discussion of the major end uses of hydrogen cyanide is presented later in this section. U.S. consumption of hydrogen cyanide grew at an average rate of 9 percent in the 1980s and was expected to maintain the same rate into the 1990s.

Hydrogen cyanide is an extremely toxic chemical in both vapor and liquid states. The threshold odor limit for most people is <5 ppm, but the odor is not noxious and often is compared to the fragrance of bitter almonds. It is highly soluble in water, alcohols, and ethers. Although HCN is a colorless liquid, iron cyanide is blue (prussian blue); and this slight contaminant usually gives HCN a bluish tint that lead to the German name *Blausäure*—blue acid. The acidity of hydrogen cyanide is weak, with a dissociation constant comparable to that of neutral amino acids. It polymerizes at pH conditions between 5 and 11, as in the presence of ammonia and water, and the polymerization can become explosively violent if confined. The change of appearance from colorless to yellowish brown is the first sign of polymerization. Disposition of HCN by combustion is considered safe provided that ventilation is adequate. Alternatively, a hydrogen cyanide solution can be detoxified by first reacting HCN with excess caustic to form sodium cyanide and subsequently reacting it with ferrous sulfate solution in a 1:1 weight ratio. Ferro-cyanide, the resulting stable complex, is relatively nontoxic.

Instruments for detection and alarm for low concentrations of hydrogen cyanide are readily available. Also a test paper for detecting HCN can be prepared by moistening filter paper with a 1:1 mixture of cupric acetate solution (3 g/liter) and benzidine acetate solution (60 vol. %). The test paper will turn blue if the HCN content is greater than 20 ppm.

Selected physical properties of hydrogen cyanide are tabulated in Table 28.19.

Manufacture

Hydrogen cyanide can be manufactured from almost any form of hydrogen, carbon, and nitrogen if sufficient energy is provided. However, only five processes are presently of economic significance. The Andrussow process is currently the principal HCN manufacturing process in the world. The BMA process is practiced by one company and exhibits a high yield and selectivity by using a complex reaction system. The Fluohmic process is of interest in locations where electricity is inexpensive. The formamide process is useful for sites with inexpensive carbon monoxide. The Sohio acrylonitrile process produces HCN as a by-product.

TABLE 28.19 Physical Properties of Hydrogen Cyanide

Molecular formula	HC≡N
Molecular weight	27.03
Melting point	−13.2°C
Boiling point (1 atm)	25.7°C
Vapor pressure	
−10°C	166 mm Hg
0°C	264 mm Hg
10°C	409 mm Hg
20°C	612 mm Hg
Density	
Gas @31°C	0.765 lb/ft^3
Liquid @20°C	42.9 lb/ft^3
Viscosity @20°C	0.201 cP
Heat-of-formation @25°C, 1 atm	
Vapor	−130.5 kJ/mol
Liquid	−105.4 kJ/mol
Heat of combustion	664 kJ/mol
Heat of polymerization	42.7 kJ/mol
Heat of fusion	8.4 kJ/mol
Latent heat	23.8 kJ/mol
Dissociation constant @25°C	7.2×10^{-10}
Conductivity, s/cm	3.3×10^{-6}
Flammability in air	
Low	5.6%
High	40%

Andrussow Process. This process, developed by Dr. L. Andrussow in Germany in the early 1930s, has been the major process for large-scale HCN manufacture for the last three decades. (See Fig. 28.30.) The reaction of ammonia with methane to form HCN is as follows:

$$NH_3 + CH_4 \rightarrow HCN + 3H_2 - 59.95 \text{ kcal/g mol}$$

The reaction is very endothermic. The Andrussow process adds additional methane and air for *in situ* combustion to heat the reactants to the 1100°C reaction temperature and to provide the heat of reaction. The overall reaction then is:

$$NH_3 + CH_4 + O_2 \xrightarrow[\text{Pt}]{1100°C} HCN + 2H_2O + H_2 + 8.5 \text{ kcal/mol}$$

In practice the reaction does not convert all of the ammonia or methane, and some carbon oxides are formed. The reaction mechanism

Fig. 28.30. Andrussow process for HCN.

actually is a sum of 13 simultaneous, sequential, and competing reactions.

A platinum/rhodium alloy is employed as catalyst for the reaction at a ratio of 9:1 as in nitric acid manufacture. Platinum alone is not stable at reaction conditions. When it is alloyed with rhodium, the catalyst life can range from 4000 to as much as 10,000 hours. The catalyst bed usually is made of layers of 40 mesh gauze made of alloy wire 0.003 inch in diameter. The spent catalyst is mostly recoverable.

The reaction takes place under fuel-rich conditions. Typical feed composition to the converter is 13–15 percent ammonia, 11–13 percent methane, and 72–76 percent air on a volumetric basis. Control of the feed composition is essential to guard against explosions as well as to maximize the yield.

The reaction usually is initiated by preheating a section of the catalyst bed with electricity or a torch. Once ignited, the reaction is self-sustained by the exothermic heat. The performance of a typical Andrussow reaction system is:

	Conversion	*Yield*	
	(mole %)	*(mole %)*	*(lb/lb HCN)*
CH_4	60–80%	60–70%	0.85–1.0
NH_3	60–70%	60–90%	0.70–1.05

The crude product at the exit of the converter consists of:

Product	*mole %*
HCN	6–12
NH_3	1.5–3
H_2	7–11
N_2	50–55
H_2O	20–25
CO	3–4
CO_2	0.3
CH_4	0.1

The heat of reaction is recovered in a waste heat boiler to generate 2 to 6 lb of steam/lb of HCN produced. The crude product then is quenched to below 350°C to avoid cracking the HCN. The unreacted ammonia is removed from the off-gas in an absorption system, and the ammonia-free off-gas then is passed through the HCN absorber, from which the flammable waste gases are burned for energy recovery. Water is used as the absorbent, with H_2SO_4 added to inhibit the polymerization of HCN. In the final step of the refining process, the bottoms of the HCN absorber are concentrated in the HCN stripper and rectifier to >99 percent hydrogen cyanide with ~0.1 percent H_2SO_4 added as a polymerization inhibitor.

Several systems have been patented for removal of the unreacted ammonia from the crude product gas. The early Andrussow processes used dilute sulfuric acid (~6%) as absorbent and produced ammonium sulfate as a by-product. However, the high cost of processing this material has made it increasingly difficult to compete in the already oversupplied fertilizer market. Unless there are special circumstances, present economics and environmental considerations favor the use of a recycle system to recover the unconverted amonia. The typical ammonia recovery systems use monoammonium phosphate at ~80°C as the absorbent to react with the unconverted ammonia in the crude product gas to form diammonium phosphate. The diammonium phosphate solution then undergoes a thermal reversal process to liberate ammonia for recycle. The regenerated monoammonium phosphate solution is reused in the absorption system. Polyhydroxyboric acid complex is an alternative absorbent for ammonia recovery that works on the same principles.

The feed gas systems are equipped with filters, and some are equipped with scrubbing systems to ensure the quality of the feed gas. Sulfur compounds in the feed are known to deactivate the catalyst. Also, unsaturated hydrocarbons will crack to coke over the catalyst screen. Reportedly, iron rust poisons the catalyst. A dual system for filtering ammonia and air separately is recommended because acidic compounds in air are known to react with ammonia to form soots and cause plugging in the feed system.

Because the converter is operated near flame conditions, the linear velocity of the reactant

should be higher than 14 ft/sec to avoid flash back. Reportedly, a reaction residence time of 10^{-8} second is adequate.[109,110] The Andrussow process has the advantage of a simple reaction system in comparison with the BMA and Fluohmic processes. However, the yield and the conversion of both methane and ammonia in the Andrussow process are the lowest among the major commercial processes. Thus, future rising costs of raw materials are expected to affect the economic attractiveness of the Andrussow process. The low HCN concentration in the crude product gas (6–7%) also requires larger equipment and higher energy consumption per pound of HCN compared with the other processes.

In 1963, Jenks and Shepard of Du Pont[111] reported that, by preheating the feed gases to 505°C (vs. 95°C conventional), the HCN concentration in the crude can be increased to 10 to 15 percent with an improved ammonia yield. Even with this improvement, the HCN concentration in the crude still is substantially below that of the BMA and Fluohmic processes.

BMA Process. This route to HCN was developed by Degussa of Germany in 1949. According to one source,[112] the reaction involves the direct combination of ammonia and methane (see first equation in this section). The reaction takes place on the inside of ceramic alumina tubes that are coated on the inside with a very thin (15 microns) layer of platinum alloyed with minor amounts of other metals (such as ruthenium and aluminum). The heat of reaction is provided outside the tubes by burning natural gas. To achieve adequate heat transfer the tubes are quite small, about $\frac{3}{4}$ inch in diameter and $6\frac{1}{2}$ feet long.

The catalyst life is typically 6500 to 10,000 hours. Reportedly, a BMA reactor of 125 metric tons/month HCN capacity is comprised of 260 reactor tubes fired to 1200 to 1300°C. The energy requirement of the reaction system is in the range of 18,000 to 20,000 Btu/lb HCN, with much of this heat recovered as steam in waste heat boilers. The performance of the BMA reaction system is reported to be:

	Conversion (mole %)	*Yield (mole %)*	*Yield (lb/lb HCN)*
CH_4	90–91%	90–93%	0.64–0.66
NH_3	83–85%	84–91%	0.69–0.75

The crude product at the exit of the reactor contains:

Product	*mole %*
HCN	20–23
NH_3	2–3
CH_4	1–2.4
N_2	0.5–1
H_2	70–72

The unreacted ammonia is removed from the crude by scrubbing with sulfuric acid for disposal as ammonium sulfate. Similarly to the Andrussow process, the HCN is separated from the crude via absorption into water with H_2SO_4 added as inhibitor against polymerization. The vent gas from the HCN absorber consists of:

Product	*Vol. %*
Hydrogen	96
CH_4	2.5
N_2 and O_2	1.5

The relatively pure hydrogen by-product is used either as a chemical feedstock or as fuel gas in the furnace. The generation rate of hydrogen is in the range of 0.19 to 0.23 lb H_2/lb HCN.

The higher conversion and yield of the BMA process have an advantage over the Andrussow process in lower production cost and small equipment size except for the reactor system. Unfortunately, the long residence time required for the reaction makes the process less attractive for large-scale production.[113] The complexity of the reactor system and the fragility of the reactor tubes are additional drawbacks.

Improvement of the heat transfer mechanism in the reaction system is an area of interest. The conventional BMA process requires heat transfer across the nonconductive reactor

tubes. U.S. Patent 1,190,922[114] introduced the use of a "pebble" heat exchanger where a circulating stream of ceramic pebbles is used as the heat transfer medium. The pebbles are heated in a combustion vessel and then pumped or fluidized to another vessel where they heat the reactants. Other mechanisms, such as rotary preheaters, also have been studied. However, none of these has found commercial application.

Fluohmic Process. This process was developed in the early 1960s by Shawinigan Chemicals. The reaction involves:

$$3NH_3 + \underset{\text{(e.g., propane)}}{\text{Hydrocarbon}} + 151\ \text{kcal/mol} \xrightarrow{1370°C} 3HCN + 7H_2$$

No catalyst is required for the reaction. The heat input to the endothermic reaction is furnished by an electrically heated fluidized bed of coke particles.

The performance of the reaction system is reported to be:

	Conversion (mole %)	*Yield (mole %)*	*Yield (lb/lb HCN)*
C_3H_8	88%	88–90%	0.60–0.62
NH_3	86%	86–90%	0.70–0.73

The crude product gas at the exit of the reactor consists of:

Product	*mole %*
HCN	25
NH_3	0.25
H_2	72
N_2	3

The HCN is separated from the off-gas for refining, and the hydrogen leaves the system as a by-product together with a small amount of nitrogen and unreacted ammonia. The by-product hydrogen stream is sufficiently pure for chemical use; for example, to manufacture ammonia if the two manufacturing facilities are within pipeline distances. Alternatively, the hydrogen can be used as fuel.

The Fluohmic process has the advantage of high HCN concentration in the crude as well as high conversion of ammonia and propane. The low ammonia concentration in the crude eliminates the need for recovery facilities. Because H_2O is absent from the system, the formation of the undesirable "azulmic" acids is precluded. However, because of the high electricity consumption (~3 kWh/lb HCN produced), the process is attractive only in locations where the cost of electricity is extremely low. Even with moderately low-cost electricity such as that on the U.S. Gulf Coast, the cost of HCN production still would be higher by this process than with the Andrussow process. Reportedly, the Fluohmic process can be used with heavy or light naphtha and is of interest in areas where methane is unavailable.

Formamide Process. Developed independently by Du Pont and BSAF, this process was used commercially in Japan and Western Europe; but since the 1950s it largely has been replaced by the Andrussow process except for BASF in Germany. The reaction involves the following steps:

(1) Methyl formate synthesis:

$$CO + CH_3OH \xrightarrow[150\text{–}3500\ \text{psi}]{100°C} \underset{\text{Methyl formate}}{HC(=O)—OCH_3}$$

(2) Formamide synthesis:

$$HC(=O)OCH_3 + NH_3 \xrightarrow[200\ \text{psi}]{40°C} \underset{\text{Formamide}}{HC(=O)NH_2} + CH_3OH$$

(3) HCN formation:[115,116]

$$HC(=O)NH_2 + 20.4\ \text{kcal/mol} \xrightarrow[\text{Acidic } Fe_2O_3]{400°C\ VAC} HCN + H_2O$$

The conversion of ammonia is 60 to 85 percent. The overall performance of the

reaction system starting from the synthesis of methyl formate is:

Product	Yield (mol %)	(lb/lb HCN)
NH_3	79–90%	0.7–0.8
CO	52–94%	1.1–2.0
CH_3OH	—	0.2–0.3

The crude product consists of 60 to 70 vol. % of HCN. The process is of interest for the *in situ* manufacture of small amounts of hydrogen cyanide. Because the transport of formamide is easier and safer than that of methane, ammonia, or HCN, the process is useful for areas where methane and ammonia are unavailable. The conversion of formamide to HCN (Reaction 3) is reported to be in the range of 92 to 95 mole %.

U.S. Patent 4,869,889 by M. B. Sherwin and Jow-Lih Su[117] covers an improved process by combining the exothermic methanol ammoxidation with the endothermic formamide decomposition.

Sohio Acrylonitrile Process. This process, developed in the late 1950s by Standard Oil Company of Ohio, produces hydrogen cyanide and acetonitrile as by-products of acrylonitrile production. Ironically, the manufacture of acrylonitrile used to be the largest consumer of HCN until the Sohio process became the principal source of acrylonitrile in the 1960s. The by-product from the Sohio process was estimated to have supplied ~20 percent of the HCN demand in 1990. The chemistry involves:

Primary reaction:

$$\underset{\text{Propylene}}{CH_2{=}CHCH_3} + NH_3 + \tfrac{3}{2}O_2 \xrightarrow{\text{Cat.}} \underset{\text{Acrylonitrile}}{CH_2{=}CHCN} + 3H_2O$$

Side reactions:

$$\underset{\text{Propylene}}{CH_2{=}CHCH_3} + \tfrac{3}{2}NH_3 + \tfrac{3}{2}O_2 \xrightarrow{\text{Cat.}} \underset{\text{Acetonitrile}}{\tfrac{3}{2}CH_3CN} + 3H_2O$$

$$\underset{\text{Propylene}}{CH_2{=}CHCH_3} + 3NH_3 + 3O_2 \xrightarrow{\text{Cat.}} \underset{\text{Hydrogen cyanide}}{3HCN} + 6H_2O$$

The conventional Sohio process produces 0.1 lb HCN/lb acrylonitrile or 0.10 to 0.13 lb HCN/lb propylene.

In the 1970s, new catalysts were developed to boost acrylonitrile yield by curtailing by-product formation. It was claimed that these catalysts produced on average ~0.06 lb HCN/lb acrylonitrile or ~0.07 lb HCN/lb propylene. Because acrylonitrile had a greater commercial value than hydrogen cyanide, this development was important economically. In the 1980s the importance of the HCN increased, and methods were developed to increase HCN production by feeding methanol to the acrylonitrile reactors.

Other Routes to HCN. The Andrussow and BMA processes are difficult to surpass on the basis of raw material cost. Methane is by far the most economical source of carbon except for coal. Ammonia is the cheapest source of nitrogen unless there is a technological breakthrough in the use of elemental nitrogen. A brief description of several processes using coal or nitrogen is presented below. However, these processes have yet to demonstrate commercial significance.

1. $N_2 + H_2 + C \xrightarrow{2000°C} HCN$, where the electric arc was generated by two carbon electrodes. Only low conversion has been reported.
2. $2CH_4 + N_2 \xrightarrow{2200°C} 2HCN + 3H_2$, which is described in U.S. Patent 2,776,872.[118]
3. Ammoniated

$$\text{coal} + NH_3 \xrightarrow{1000\text{–}1300°C} HCN + H_2,$$

developed by the U.S. Bureau of Mines (U.S. Patent 3,501,267).[119] The conversion of ammonia was close to 100 percent, and the yields obtained were 0.9 to 1.4 lb HCN/lb NH_3 and 0.04 to 0.20 lb

HCN/lb coal. The crude product composition was 11 percent HCN. However, the data reported were obtained under a partial helium atmosphere, and performance under realistic conditions may be difficult. The process is of possible interest in coal mining states where the cost of carbon from coal is a fraction of that from methane. More work is required to commercialize the technology.

In special circumstances where the availability of hydrocarbons is a logical constraint, several alternative processes can be considered, especially for small-scale production of HCN:

1. $CH_3OH + O_2 + NH_3 \rightarrow HCN + 3H_2O$, which has been reviewed by Stanford Research Institute. Early patents reported only 3 mole % of HCN in the crude product. However, a recent patent showed improvement of the process to approach 70 percent of the Andrussow process.[120]

2. $$\underset{\text{Toluene}}{CH_3C_6H_5} + NH_3 \xrightarrow[1000^\circ C\ N_2]{Pt,\ PtO_2Cat} \underset{\text{Benzene}}{C_6H_6} + HCN + 2H_2$$

   where the yields obtained were 100 percent conversion of toluene to benzene and 77 percent yield of HCN.[119]
3. $2NaCN + H_2SO_4 \rightarrow Na_2SO_4 + 2HCN$. This reaction is used for preparation of small quantities of hydrogen cyanide.

In addition, Japanese Patent 13459[121] reported noncatalytic reaction $2CH_4 + NH_3 + \frac{1}{2}O_2 \rightarrow HCN + 5H_2 + CO$. There has been no commercial application of this process.

Economics and Marketing

The U.S. HCN capacity in 1990 was reported to be 1.47 billion lb/year, and business forecasts projected 75 percent utilization of capacity in 1990. Major U.S. producers of HCN and the capacities of their manufacturing facilities are summarized in Table 28.20.

TABLE 28.20 U.S. Producers of Hydrogen Cyanide

Company	*Annual Capacity as of June 1990 (millions of pounds)*
American Cyanamid Company (B)	40
BP Chemicals (B)	110
Ciba-Geigy Corporation	90
Degussa Corporation	53
Dow Chemical U.S.A.	20
E. I. Du Pont de Nemours & Co. Inc.	900
Monsanto Company (B)	60
Röhm and Haas Company	200
Sterling Chemical (B)	90
Total	1563

(B)—By-product from acrylonitrile manufacture.
Source: *Chemical Economics Handbook*, 664,5020E.

End Uses. Hydrogen cyanide is the main raw material for the manufacture of numerous products, including:

- Adiponitrile

  $(H_2NCH_2CH_2CH_2CH_2CH_2CH_2NH_2)$,

  made by reacting HCN with butadiene. This is in turn used for the manufacture of hexamethylene diamine, which finds its way into the manufacture of nylon. Adiponitrile manufacture was the largest single use of HCN in 1991 (Table 28.21).
- Methyl methacrylate

  $(CH_2{=}C(CH_3)COOCH_3)$,

  made by reacting HCN with acetone, sulfuric acid, and methanol. Methyl methacrylate is used in making clear acrylics, which are important plastics (Chapter 19).
- Sodium cyanide, from the reaction of HCN with sodium hydroxide. This product is useful in the recovery of gold from low-grade ore.

TABLE 28.21 U.S. Approximate Consumption of Hydrogen Cyanide, 1991

	Millions of Pounds of HCN
Adiponitrile	500
Acetone cyanohydrin	430
Sodium cyanide	150
Methionine	49
Cyanuric chloride	45
Chelating agents	25
Nitrilotriacetic acid	20
Other	45

Source: SRI, "Hydrogen Cyanide," Feb. 1989.[113]

- Cyanuric chloride ($C_3N_3Cl_3$), which is an intermediate for the manufacture of pharmaceuticals, explosives, herbicides, etc.
- Chelating agents, made by reacting HCN with formaldehyde and amines to yield amino carboxylic acids, which then are saponified to produce the acid salts. The more common acids are NTA, EDTA, HEDTA, DTPA, and others. Their major uses are in the manufacture of soaps, water treating agents, and various cleaning agents.
- Methionine, an additive in poultry feeds.
- Acetone cyanohydrin.

Table 28.21 shows the approximate U.S. consumption of HCN in 1991. The list price for hydrogen cyanide in tank cars f.o.b. was 60 cents/lb in 1991.[46]

Licensers are:[47]

Du Pont
Mitsubishi Gas
Mitsui Toatsu
Polimex-Cekop
Röhm GmbH

ANILINE

Aniline or aminobenzene (benzamine) was produced first in 1826 by Unverdorben, by the dry distillation of indigo, the oldest known vat dye. Fritsche also obtained this material from indigo by heating it with potash, and he named it aniline. Hofmann obtained it by reduction of nitrobenzene in 1843, proving the structure. Aniline is a colorless, oily, flammable liquid, which is slightly soluble in cold water and infinitely soluble in alcohol and ether. It is highly toxic with a threshold limit value of 5 ppm by volume. Its physical properties are summarized in Table 28.22, and its vapor pressure characteristics in Table 28.23.

TABLE 28.22 Physical Properties of Aniline

Property	*Value*
Boiling point, °C	
101.3 kPa (760 mm Hg)	184.4
4.4 kPa (33 mm Hg)	92
1.2 kPa (9 mm Hg)	71
Melting point, °C	−6.15
Density, *d*	
at 20/4°C	1.02173
at 20/20°C	1.022
Viscosity at 20°C, mPa·s (=cP)	4.423–4.435
Dissociation constant, *pK*	
at 20°C	4.60
at 60°C	8.88
Enthalpy of dissociation, kJ/mol (kcal/mol)	21.7 (5.19)
Heat of combustion, kJ/mol (kcal/mol)	3389.72 (810.55)
Specific heat, 20–25°C	0.518
Latent heat of vaporization, J/g (cal/g)	476.3 (113.9)
Flash point (closed-cup), °C	76

TABLE 28.23 Vapor Pressure of Aniline

Temperature °C	*Vapor Pressure kPa*	*Temperature °C*	*Vapor Pressure kPa*
175	79.99	139	27.67
162	53.33	119	13.33
151	39.99	102	6.67

Aniline is produced from nitrobenzene, which, in turn, is produced by reacting nitric acid with benzene at 50°C in the presence of sulfuric acid. The reaction is postulated to proceed as follows:

$$HONO_2 + H^+ \rightarrow H_2ONO_2{}^+$$

$$H_2ONO_2{}^+ \rightleftarrows H_2O + NO_2{}^+ \text{ (nitronium ion)}$$

Benzene Nitrobenzene

In concentrated sulfuric acid, HNO_3 is completely converted to NO_2^+, thus favoring the nitration reaction.

Impurities in the nitrobenzene result from impurities in the benzene such as nitrocresols, and other nitration products such as nitrous acid, nitrogen oxides, dinitrobenzene, nitrotoluene, carbolic acid, and nitrophenols. More details on the production of nitrobenzene are reviewed in Chapter 23.

Aniline is produced by hydrogenation of the nitrobenzene amine group to an amino group. A catalyst is used such as copper/silica, nickel, or platinum/palladium. The reaction can be carried out in either the vapor or the liquid phase at 250 to 300°C and 10 to 25 atmospheres.

Nitrobenzene Aniline

Yields from nitrobenzene are reported to be greater than 99 percent of theoretical.

Some aniline still is produced by reducing nitrobenzene with iron in the presence of ferrous chloride and/or hydrochloric acid.

The Scientific Design, Inc. process for aminolysis of phenol is shown in Fig. 28.31.[122] In this process, a mixed-vapor feed of phenol and ammonia is preheated and passed over a fixed catalyst bed. A unique SD-developed catalyst is used. The reactor effluent is partially condensed and then sent to an ammonia recovery still. The recovered ammonia is recycled, and the ammonia recovery still bottoms are stripped of ammonia, dried, and distilled. Nearly stoichiometric yields are said to be obtained, and the aniline is of high purity. An azeotrope of aniline and phenol from the purification still is recycled to the reactor, and a long catalyst life is claimed. A 30,000 metric ton/year plant has been operated by Mitsui Petrochemical Industries, Ltd. since 1970. In 1882, U.S. Steel put a 90,000 metric ton/year plant onstream. This process is favorable where low cost phenol is available, and when high purity aniline is desired. Capital costs are low because the nitration of benzene is avoided. Waste disposal problems are minimal.

Reduction of nitrobenzene also can be carried out in the liquid phase with H_2S, but this is not a commercial process at present.

Fig. 28.31. Aniline by aminolysis of phenol. (*"Aniline,"* Hydrocarbon Processing, Petrochemical Handbook *issue, p. 124, Nov. 1985. Copyright by Gulf Publishing Company. By permission.*)

A fixed bed of catalyst is used in the Lonaz process, which has been installed at the First Chemical Corp. plant in Pascagoula, Mississippi.[123] The plant design rate is 250,000 lb/year. Yield is about 85 percent, or 1.35 lb of nitrobenzene/lb of aniline.

Aminolysis of chlorobenzene also will produce aniline although it probably is not used commercially.

$$\underset{\text{Chlorobenzene}}{C_6H_5Cl} + 2NH_3 \rightarrow \underset{\text{Aniline}}{C_6H_5NH_2} + NH_4Cl$$

For the aminolysis of chlorobenzene, 3 moles of ammonia usually are used at an operating temperature of 180 to 200°C and a pressure greater than the vapor pressure of the reactants. This reaction is carried out in the liquid phase with copper compounds as catalysts.

Aniline is used primarily for the production of 4,4′-methylenebis(phenylisocyanate) which consumes about 75 percent of the aniline produced. Other products include rubber chemicals (13%), dyestuffs (3%), synthetic fibers (2%), pharmaceuticals (1%), and agricultural chemicals (6%). U.S. producers of aniline and their capacities are:

	Million pounds per year
Aristech Chemical Corp.	200
E. I. Du Pont de Nemours and Co.	260
Mobay Chemical Corp.	40
Rubicon Chemicals, Inc.	360
First Chemical Corp.	250

The demand in 1988 was 1 billion lb. Growth in the 1980s averaged 4.2 percent, but most came in the last half of the decade when growth averaged 8.5 percent/year.

The list price for aniline in 1991 was 50 cents/lb in tank cars, f.o.b.[46] Process licensers for aniline in 1990 were:[47]

American Cyanamid
Du Pont
Lonza
Josef Meissner
Mitsui Toatsu
Petroquisa/Pronor
Scientific Design
Sumitomo Chemicals
Tolochimie

OTHER COMPOUNDS

Several other nitrogen compounds are of commercially important, including:

- Hexamethylenediamine (Chapter 21), used primarily in the manufacture of nylon.
- Ethanolamines (Chapter 22), used in the manufacture of detergents and as scrubbing agents for the removal of acidic compounds from gases.
- Acrylonitrile (Chapters 18, 19, 21, and 22), used in the manufacture of acrylic and modacrylic fibers, plastics and resins, and nitrile rubbers or elastomers.
- Dimethylformamide, a versatile solvent for organic and inorganic compounds and an important reaction medium for ionic and nonionic compounds.
- Dimethylacetamide, an important industrial solvent for polyacrylonitrile, vinyl resins, cellulose derivatives, styrene polymers, and linear polyesters.
- Isocyanates, important materials in the production of foams, resins, and rubbers.

REFERENCES

1. *Chem. Eng. News*, p. 16 (Aug. 27, 1979).
2. Harris, G. T., and Harre, F. A., "World Fertilizer Situation and Outlook, 1978–85—and Trends 1979," TVA, Muscle Shoals, AL, 1979.
3. SRI International, *Chemical Economics Handbook*, 1988.
4. *Hydrocarbon Proc.*, Section 2 (Oct. 1979).
5. Updegraff. N. C., and Mayland, B. J., *Petrol. Refiner*, **33** (12), 156–159 (Dec. 1954).
6. *Chem. Eng.*, pp. 57–59 (Oct. 8, 1979).
7. U.S. Patent 4,153,673.

8. U.S. Patent 4,162,290.
9. U.S. Patent 3,106,457.
10. Hougen, O. A., et al., *Chemical Process Principles*, Part II, p. 1107, John Wiley & Sons, New York, 1959.
11. *Catalyst Handbook*, p. 79, Katalco Corp., Oak Brook, Ill.
12. U.S. Patent 3,827,987.
13. *Chem. Eng.*, pp. 115–128 (Dec. 18, 1978).
14. A.C.S. Synthetic Fuels Symposium, Sept. 12–15, 1955.
15. Am. Petrol. Inst., "Selected Values of Properties of Hydrocarbons," Research Project 44, Nat. Bureau of Standards.
16. *Chem. Eng.*, pp. 69–71 (Jan. 30, 1978).
17. Lombard, J. F., *Hydrocarbon Proc.*, **48** (8), 111 (Aug. 1969).
18. Lundberg, W. C., *Chem. Eng. Prog.*, p. 81 (June 1979).
19. Slack, A. V., and James, G. R., *Ammonia*, Part III, pp. 123–153.
20. *Chem. Eng. Prog.*, **69** No. (2), 67–70 (Feb. 1973).
21. AICHE, "Amine Guard Systems in Hydrogen Production," 86th National Meeting, Houston, TX, Apr. 1–5, 1979.
22. U.S. Patent 3,352,631.
23. U.S. Patent 3,347,621.
24. *Chem. Eng.*, pp. 88–89 (Dec. 3, 1979).
25. *Hydrocarbon Proc.*, pp. 145–151 (Apr. 1978).
26. U.S. Patent 2,863,527.
27. *Chem. Eng.*, **67** (19), 166–169 (Sept. 19, 1960).
28. *Catalyst Handbook*, p. 119, Springer-Verlag, New York.
29. "Removing Carbon Monoxide from Ammonia Synthesis Gas," *Ind. Eng. Chem.*, **53** (8), 645 (Aug. 1961).
30. Grotz, B. J., *Hydrocarbon Proc.*, **46** (4), 197 (1967).
31. U.S. Patent 3,584,998.
32. Kellogg Ammonia License Bulletin, Reprint 1990.
33. "Compression Systems for Ammonia Plants," *Chem. Eng. Prog.*, p. 88 (July, 1979).
34. Quartulli et al., "Best Pressure for Ammonia Plants," *Hydrocarbon Proc.*, **47** (11), 153–161 (1968).
35. Nielsen Andres, *An Investigation of Promoted Iron Catalysts for the Synthesis of Ammonia*, 3rd ed., Jul. Gjellerup Forlag, Copenhagen, 1956.
36. *Catalyst Reviews*, pp. 1–25, (Apr. 9, 1970).
37. Ulmanns, *Encyklopadie de Technischen Chemie*, Urban and Schwarzenberg, Munchen and Berlin, 1953.
38. "Viscosity–Temperature Correlation for Liquids," *Chem. Eng.*, p. 83 (July 16, 1979).
39. U.S. Patent 2,894,821.
40. U.S. Patent 3,992,328.
41. *Chem. Eng.*, p. 63 (July 16, 1979).
42. U.S. Patent 3,310,376.
43. Slack, A. V., and James, G. R., *Ammonia*, Part III, pp. 291 and 387.
44. Kellogg Bulletin, "Large Scale Production of UAN," 1990.
45. Bechtel License Bulletin, "Ammonia," 1991.
46. *Chemical Marketing Reporter* (Apr. 29, 1991).
47. *ECN Process Review* (Oct./Nov. 1990).
48. Sauchelli, "Fertilizer Nitrogen," ACS Monograph Series 161.
49. Thompson, R. *The Modern Inorganic Chemical Industry*, The Chemical Society, Burlington House, London, 1977.
50. *Hydrocarbon Proc.*, Petrochemical Handbook Issue, p. 119 (Nov. 1983).
51. "Nitric Acid, SABAR," *Hydrocarbon Proc.*, p. 106 (Nov. 1989).
52. Spratt, D. A., *Proc. Fertilizer Soc. of England* (Mar. 25, 1958).
53. "Catalytic Processes in Nitric Acid Manufacture," *Fertilizer Society Proceedings*, London, 1978.
54. U.S. Patent 4,957,720.
55. *Chem. Eng.*, **67** (8), 94 (1960).
56. U.S. Patent 3,202,481.
57. U.S. Patent 2,760,845.
58. Powell, R., "Nitric Acid Technology—Recent Developments," Noyes Development Corp., Park Ridge, NJ, 1969.
59. *Chem. Eng. News*, **44** (27), 13 (July 4, 1966).
60. SRI International, *Chemical Economics Handbook*, 1987.
61. Stafford, J. D., Samuels, W. E., and Croysdale, L. G., "Ammonium Nitrate Plant Safety," Am. Inst. of Chem. Eng., Sept. 27–29, 1965.
62. *Chem. Week*, **101** (2), 75 (July 8, 1967).
63. *Hydrocarbon Proc.* (1987).
64. Franz, R., and Applegath, F., *Chem. Eng.*, **67** (25), 71 (1960).

65. U.S. Patent 3,232,984.
66. Bechtel Corporation Bulletin, "Urea," 1991.
67. U.S. Patent 3,107,149.
68. "Urea Process Technology," p. 139, Noyes Development Corp., Park Ridge, NJ.
69. U.S. Patent 3,236,888.
70. U.S. Patent 3,200,148.
71. U.S. Patent 3,301,897.
72. U.S. Patent 2,527,315.
73. Harri, E. A., "The Outlook for Nitrogen Fertilizers," TVA, Muscle Shoals, AL, Forest Fertilizer Conference, Union, Washington, Sept. 25–27, 1979.
74. *Hydrocarbon Proc.*, **48** (11), 209 (Nov. 1969).
75. U.S. Patent 3,498,982.
76. Takakuwa, Y., et al., "Development of a Process for the Purification of Crude Melamine Manufactured from Urea," *Int. Chem. Engineering*, **19** (4), 624 (Oct. 1979).
77. SRI, "Directory of Chem. Products," Jan. 1990.
78.
79. Novella and Una, *An. Quim.*, **65**, 699–708 (1969).
80. Egly, R. S., and Smith, E. F., CEP44, No. 5, *Trans. Am. Inst. Chem. Engrs.*, pp. 387–398 (1948).
81. U.S. Patent 1,799,722.
82. U.S. Patent 3,387,032.
83. E. German Patent 108,275.
84. Japan Patent 1623 (1952).
85. West German Patent DT-2916-060.
86. Serban, S., *Rev. Chim. (Bucharest)*, **14** (8), 451–454 (1963).
87. Issoire, J., and Long, C. V., *Bull. Soc. Chem. France*, pp. 2004–2012 (1960).
88. U.S. Patent 3,720,715.
89. Ulmanns, *Encyklopadie de Technischen Chemie*, 3rd ed., Vol. 12, pp. 421–425, Munich and Berlin, Urban and Schwarzenberg, 1960.
90. "Methylamines," *Hydrocarbon Proc.*, pp. 114, 115 (Nov. 1983).
91. U.S. Patent 3,444,203.
92. DBP 1,019,313 (Old German Patent).
93. French Patent 762194.
94. Schnell Publishing Co., Inc., *Chemical Profile* (July 1, 1979).
95. SRI International, *Chemical Economics Handbook*, "Alkylamines," 1988.
96. Sitig, M., "Amines, Nitriles and Isocyanate Processes and Products," Noyes Development Corp., Park Ridge, NJ, 1969.
97. SRI, *Chemical Economics Handbook*, "Ethyleneamines," 1989.
98. Astle, M. J., *Industrial Organic Nitrogen Compounds*, pp. 12–13, 1961.
99. Meissner, F., Schwiedessen, E., and Othmer, D. F., *Ind. Eng. Chem.*, **46**, 724 (1954).
100. U.S. Patent 3,538,199.
101. Czechoslovakian Patent 250012.
102. *Chemical Economics Handbook*, 658, 5033X.
103. SRI International, "Directory of Chemical Producers," 1990.
104. Schmidt, *Hydrazine and Its Derivatives*, p. 485, John Wiley & Sons, New York, 1984.
105. *Chem. Eng.*, (1), 120–123 (July 14, 1958).
106. Japanese Patent Application SHO-62-83308.
107. Powell, R., "Hydrazine Manufacturing Processes," Noyes Development Corp., Park Ridge, NJ, 1968.
108. Schmidt, *Hydrazine and Its Derivatives*, p. 514, John Wiley & Sons, New York, 1984.
109. Andrussow, L., *Angew Chem.*, **48**, 593 (1935).
110. Kirk and Othmer, *Encyclopedia of Chemical Technology*, Vol. 6, p. 577, John Wiley & Sons, Inc., New York, 1980.
111. U.S. Patent 3,104,940.
112. *Hydrocarbon Proc.* **46** (11), 189 (Nov. 1967).
113. SRI International, "Hydrogen Cyanide," Feb. 1989.
114. U.S. Patent 1,190,922.
115. German Patent 944,547.
116. German Patent 1,209,501.
117. U.S. Patent 4,869,889.
118. U.S. Patent 2,776,872.
119. U.S. Patent 3,501,267.
120. Sedricks, W., "Hydrocan Cyanide from Methanol," SRI publication.

121. Japanese Patent 13459.
122. "Aniline," *Hydrocarbon Proc.*, p. 124 (Nov. 1985).
123. *Hydrocarbon Proc.*, Petrochemical Handbook Issue, Nov. 1979.
124. SRI, *Chemical Economics Handbook*, "Ammonia," p. 756.6001c.
125. *International Critical Tables*, Vol. III, pp. 304, 305, 309, McGraw-Hill Book Co., Inc., New York, 1928.
126. *International Critical Tables of Numerical Data, Physics, Chemistry, and Technology*, Vol. IV, p. 255, McGraw-Hill Book Co., Inc., New York, 1928.
127. Perry, R. H., and Chilton, C. H. (Eds.), *Chemical Engineers' Handbook*, 5th ed., McGraw-Hill Book Co., Inc., New York, 1973.
128. Schnell Publishing Co., Inc., *Chemical Economics Handbook*, 611,5030.
129. *SRI International, Chemical Economics Handbook*, "Ammonium Nitrate," 1988.
130. SRI International, *Chemical Economics Handbook*, "Nitrogen Products—Urea," July 1988.

29

The Agrochemical Industry

Gustave K. Kohn* and Don R. Baker**

INTRODUCTION

Scope of the Chapter

This chapter deals with the chemicals used in agriculture mainly to protect, preserve, and improve crop yields. The term "agrochemical" is used broadly. Much agrochemical research and some advanced development is directed toward the introduction of genes that may provide disease, insect, or viral resistance into plants or other organisms. Further progress is being made in improving the protein, fat, or carbohydrate composition of the plant itself. Microorganisms are being propagated and currently marketed that are insecticidal (e.g., *Bacillus thuringiensis*); fungi that are herbicidal, bacteria that are fungicidal, nematodes that are widely biocidal, and so on, are all products or candidate products for use in agriculture.

This chapter must deal with all of the above subject matter although it is recognized that the *classical pesticides are still the dominant products for sale* by the agrochemical industry. On a molcular basis, genes, enzymes, toxins, peptides, and viruses—whatever their source—are chemicals. If they are employed or targeted for use in agriculture, they are regarded as agrochemicals.

Arbitrarily excluded are those substances that serve as fundamental nutrients, which are treated in Chapters 11 and 28 on fertilizer and nitrogen technology, respectively. Nevertheless, it is the current practice of the farmer, particularly in advanced American agriculture, to integrate nutritional and plant-protection application schedules, and even provide single formulations that include both fertilizers and pesticides. Further, plant nutrition at this stage of scientific sophistication is far more complex than the older and conventional N–P–K applications. Many chemicals that accelerate plant growth act as hormonal agents, modifying plant metabolic processes at some stage of development. Because these substances are manufactured and marketed by the agrochemical industry, they are included as subject matter here. Also included in this chapter are chemicals that are significant in public health.

*Consultant, formerly with Sandoz Crop Protection Corp, UNIDO, and Chevron Chemical.
**ICI Americas Inc.

Many organisms are vectors in the dissemination of human and animal disease.

Because products of the pesticide industry control the insect, the rodent, the mollusk, and so forth (the vectors), they often are the most effective and sometimes the only practical means for controlling some of the most serious health problems of humankind, especially, but not exclusively, in the underdeveloped countries. Thus, chemicals that control such vectors also are included in the subject matter of this chapter. In this area, and in the area of animal health products, Chapter 25 (on the pharmaceutical industry) overlaps with this chapter to some extent.

Finally, the products of the agrochemical industry are employed in urban, suburban, and rural areas in home and garden for a variety of useful and aesthetic purposes. Moreover, industry employs herbicides, algicides, fungicides, and bactericides, and the railroads utilize herbicides to keep right-of-way free of vegetation.

The Language of the Pesticide Industry

There is a need for a generic term to cover the diversity of functional applications in this industry. The term "insecticide" has been commonly applied to all control agents. The Federal Insecticide Act of 1910 provided an extension in meaning to include both insecticides and fungicides. The Federal Insecticide, Fungicide, and Rodenticide Act (1947) uses "economic poison" for all these groups, and embraces herbicides as well. The term "pesticide" now is used officially to include all toxic chemicals, whether used against insects, fungi, weeds, rodents, or other organisms. Agricultural disinfectants and animal health products in many instances are also included under the term "pesticides."

"Plant protection agents" is another generic term. A common practice within the industry is to add the suffix "cide" to the group or biological unit under consideration. The term "insecticide" has both generic and specific meaning that depends upon the context. Obvious in meaning, also, are such frequently used words as fungicide, bactericide, miticide or acaricide, ovicide, herbicide, rodenticide, muscicide, molluscicide, and algicide. The practice of utilizing the "cide" suffix reflects upon the state of agricultural technology as it was practiced during the first half of the twentieth century. Whereas the objective in most cases, until recently, was to kill or attempt to eradicate the offending pest, many of the newer programs reflect more complex approaches. For that reason the term "agrochemical" is more commonly used now.

Hence the words attractants, repellents, antifeeding compounds, hormones (plant and insect growth control agents), defoliants, desiccants, and so on, all describe functional chemical agents that are within the purview of the agrochemical or plant protection industry. There are many who believe these chemicals may become more significant in the agriculture of the future.

History

It is probable that farmers' treatment of crops with foreign substances dates back into prehistory. The Bible abounds with references to insect depredations, plant diseases, and some basic agricultural principles such as periodic withholding of land in the fallow state. Homer speaks of "pest-averting sulfur."

More recently, in the nineteenth century, there was a great increase in the application of foreign chemicals to agriculture. Discovered or, more precisely, rediscovered was the usefulness of sulfur, lime sulfur (calcium polysulfides), and Bordeaux mixture (basic copper sulfates). With the exception of formaldehyde, inorganic chemicals provided farmers with their major weapons.

The earliest of the organic compounds generally were chemicals derived from natural products or crude mixtures of chemicals in states of very elementary refinement. Ground-up extracts of plant tissue were useful in the control of insects. Such extracts were employed in agriculture in many cases before the chemist had elucidated the structure or synthesized the molecule responsible for biological activity. These extracts included the pyrethroids, rotenoids, and nicotinoids, which continue to be

derived in large part from plant extracts. Crude petroleum fractions were recognized for their effectiveness in the control of mites, scale, and various fungi, as well as for their phytopathological properties.

Although a few synthetic organics were already known, the great revolution in the use of organic chemicals in agriculture roughly coincides with the period of the onset of World War II. The more important of these discoveries were DDT (Mueller—1939), 2,4-D (Jones patent—1945), benzenehexachloride (ICI and French development—ca. 1940), and the organic phosphate esters (Schrader—begun in the late 1930s, revealed in the forties). These new chemicals were so enormously more potent than their predecessors in their biological activity (frequently by orders of magnitude) that they very rapidly displaced almost all of the chemicals previously employed. The classical chemicals of today, some discovered in the 1950s and 1960s, are predominantly extensions of this almost revolutionary transition from inorganics to synthetic organics that dates from the period of World War II.

These trends are vividly illustrated in Fig. 29.1. As late as 1945, inorganic chemicals accounted for almost 75 percent of all pesticide sales, with oil sprays and natural products accounting for most of the remainder. The rapid reduction in the sales of inorganics was virtually complete by the early 1950s, and oil sprays and natural products reached a quantitatively low but constant level at approximately the same period.

It is fair to say that within the United States the agrochemical industry since the late 1950s, has been largely dominated by synthetic organic chemicals. The inorganics, oil sprays, continue to share a small portion of the total pesitide market. On the other hand, natural products, including five organisms, are increasing as products and projects for current research and development.

Federal government statistics on the supply of pesticides in the United States emphasize the dominance of the synthetic organics. Many of the botanicals are derived from imports into the United States of plant parts or their extracts. In contrast to the supply and

Fig. 29.1. Impact of synthetic organic chemicals on pesticide sales. (This chart is based on estimates for U.S. industry derived from data collected from one major chemical company (Chevron Chemical Co.), which was a manufacturer and marketer of organic and inorganic pesticides and oil sprays.) In the 1970s oil ceased to be a low-cost commodity, and this fact, combined with toxicological considerations, further reduced its percentage impact on pesticide sales. In this figure metalloorganics (maneb, zineb, etc.) are included in the organic category.

sales of chemicals from the standpoint of the derivation, it is interesting to examine trends according to the biological function of the pesticide.

The data in Table 29.1 show the enormous growth of the agrochemical industry (see also Fig. 29.2). Although fungicide use in the United States presently is the smallest of the three main classifications of pesticides, it must be remembered that in tropical and high-moisture areas the problem of plant disease is especially serious. High labor costs are largely responsible for the extremely rapid growth rate of herbicide production. Because herbicides diminish the need for cultivation (hoe and more modern mechanical weeding) in advanced agricultural countries, they will

TABLE 29.1 U.S. Production of Synthetic Pesticides (lb × 10^6)*

Year	*Insecticides*	*Fungicides*	*Herbicides*
1945	43.0	—	—
1950	177.3	13[a]	1.1
1955	331.1	34[a]	70.6
1960	351.0	44[a]	102.5
1965	421.6	55.0	262.9
1970	358.9	89.6	400.6
1975	501.0	114.7	788.0
1980	34.0	107.2	798.9
1985	245.8	109.0	755.8
1988[b]	222.9	109.6	701.8

[a]Figures adjusted to reflect production for agricultural uses.
[b]Latest year for which data were available.
Source: Data from various reports of Stanford Research Institute, USDA, and private company files.
*World Production figures increasing more rapidly, but in different proportions.

continue to find much use in the future. On the other hand, ecological pressures and the halting of herbicide use for military purposes did lower the production figures of certain insecticides (DDT and chlorinated hydrocarbons) and certain herbicides (2,4,5-T and 2,4-D) for 1970 and succeeding years.

Although the United States continues to provide the largest *national* market for agrochemicals, the intensity of such usage based on either agricultural acreage or tons of product per acre is higher in certain other countries in the world (Japan is a notable example).

The worldwide consumption of agrochemicals is now roughly double that of the United States. There is little doubt that this relationship (that is, worldwide pesticide usage in relation to U.S. consumption) will markedly increase, as will world population and its consequential requirement for greater quantities of food. Pesticide production and the pesticide industry are inextricably related to the Malthusian paradox.

CHARACTERISTICS OF THE AGROCHEMICAL INDUSTRY

Among the distinguishing characteristics of the agrochemical industry are (1) the multiplicity of chemical agents employed, (2) a

Fig. 29.2. Pilot plant view for production of field trial compounds. (*Courtesy of ICI Americas, Inc.*).

limited price range (and hence limited chemical complexity governed primarily by the economics of agricultural production), (3) a fairly rapid obsolesence of the chemicals so used, and (4) a high degree of government regulation for the production, application, shipment, and use of agrochemicals.

Government Regulation

The first state laws on insecticides were enacted in 1900 to establish standards of purity for the arsenical Paris green (which is no longer in use in agriculture in the United States). Gradually these laws were extended to cover a wide list of inorganic compounds and plant extracts, many of them extremely toxic to humans. Included in this group are such compounds as arsenic combined with copper, lead, and calcium; phosphorus pastes for ants and roaches; strychnine in rodent baits; thallium in ant and rodent baits; and selenium for plant-feeding mites. Mercury, both as a corrosive sublimate and as calomel, was used as an insect repellent and later as a seed disinfectant. Sodium fluoride was a common ant poison, and sodium cyanide, calcium cyanide, and HCN itself were general fumigants. Nicotine sulfate was used generally in the garden and on the farm. These compounds, among the most toxic of any known at that time, were widely marketed without supervision under any of the early state laws. There was no provision for public health, either in regulating the amounts applied or regarding the possible danger of minute amounts (residues) remaining on the marketed produce. The need to protect the applicator, farmer, or laborer and the public against the dangerous qualities of the insecticides, or their residues on crops, provided the motives for all the legislation that followed.

In response to reports from England of the poisoning of people eating apples grown in the United States, the Federal Food and Drug Administration in 1927 banned any fruit involved in interstate commerce if it possessed more the 3.57 ppm equivalent of arsenic trioxide. This landmark legislation established the precedent for what is commonly known today as a residue tolerance. Every agrochemical registered for sale today has such a residue tolerance, which quantitatively varies from zero to parts per billion and parts per million in the product to be consumed.

Government concern first was related to standardization of the manufactured chemical and protection of the farmer in relation to the product that he or she purchased. This was then extended to the handling of the chemical in interstate commerce, and then to protection of the consumer of raw agricultural products (apples, corn, and lettuce, for example); and in other legislation to protection of the consumer of finished goods (canned juice, margarine, cereal food, meat, and milk, for example). Included in this legislation were provisions that protect the shipper of the chemical, the applicator of the chemical, and all personnel proximal to the application of the chemical. Legislation now regulates chemicals applied to crops or foods as protective agents—pesticides, emulsifiers, solvents, packaging materials (wax, container materials, plasticizers, antioxidants, etc.). Table 29.2 summarizes many aspects of government regulation of the pesticide industry.

Toward the end of the 1960s, a new area of concern arose: the effect of the manufacture and application of pesticide chemicals *on the environment* was recognized. Concern for the environment was the subject matter of Rachel Carson's brilliant polemic, *Silent Spring*, which was published in 1962. This new focus led in 1970 to the establishment of the Environmental Protection Agency (EPA), which was given authority to regulate virtually all aspects of agrochemical manufacture and use in the United States.

Since its inception, the principal objectives of the EPA's agrochemical activities have been (1) to establish procedures that ensure that new pesticides will not pose unreasonable risks to human health and the environment, and (2) to terminate the use of those previously registered pesticides that exceed certain risk criteria.

The approach to the first of these objectives, regarding new pesticide products, has been to give careful scrutiny to a comprehensive and

TABLE 29.2 Development of Government Regulation of the Agrochemical Industry in the United States

Date	*Legislation*	*Content*
1900	State laws on product purity (Paris green)	Quality control for protection of user of chemical.
1910	Federal Insecticide Act	First specific pesticide regulation; standards for insecticides and fungicides moving in interstate commerce.
1938	Federal Food, Drug, and Cosmetics Act	Protection of public against contamination of foods, drugs, etc.
1947	Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA)	Regulation of interstate shipment of pesticides. Safety standards for the handling of chemicals in interstate commerce. Required federal *registration* for labeling.
1954	Miller Pesticide Amendment to FIFRA (Public Law 518)	Requires setting of residue tolerances for pesticides on raw agricultural product; protects consumer.
1958	Food Additives Amendment (contains Delaney Amendment)	Protects public against contamination of processed agricultural products and foods from pesticide and other contamination; protects consumer; prohibits direct food additives that cause cancer at any dosage level of test protocol.
1947–66	State enactments embodying FIFRA and food additives legislation	By 1966, 49 states and Puerto Rico had supplemented protection within state boundaries in varying degree for above federal legislation.
1970–71	Environmental Protection Agency (EPA) established by executive fiat.	EPA assumed regulatory responsibility for pesticides (previously held jointly by USDA and EPA).
1970–80	Occupational Safety and Health Act, OSHA. Enacted into law Dec. 1970; grew as an effective agency throughout decade.	Regulates to minimize hazards associated with manufacture of all chemicals including all pesticides.
1972	Federal Environmental Pest Control Act of 1972—Amendments to FIFRA.	FIFRA strengthened to comprehensive regulatory statute. Grants EPA authority to conduct extensive pre-market review of pesticide safety and efficacy data. Emphasis on human health and environmental effects of pesticides. Provides statutory authority for cancellation of registration for pesticides that pose risk of "unreasonable adverse effect."
1975	EPA established RPAR process by regulation.	Rebuttable Presumption Against Registration established as key mechanism for study of presumed risk of old pesticides.
1975–80	EPA proposed "Guidelines for Registering Pesticides."	Detailed testing requirements imposed on new pesticides for health and environmental effects.
1976	TSCA Congressional Act (Toxic Substances Control Act)	Regulation of toxic substances including intermediates, solvents, all chemical compositions. Pre-manufacturing notification required for all new substances.
1978	Federal Pesticide Act of 1978	Provides for "Conditional" registration and "Generic" registration (reregistration) of pesticides.
1981	Comprehensive Enviromental Response Compensation and Liability Act.	For cleanup of toxic wastes, spills, and dumps.
1985	California passed the Pesticide Contamination Prevention Act.	To reduce groundwater contamination.
1986	OSHA developed its Hazard Communication Standard.	Requires employers to provide Material Safety Data Sheets to employees working with or exposed to hazardous materials.

TABLE 29.2 (continued)

Date	*Legislation*	*Content*
	Safe Drinking Water and Toxic Enforcement Act of 1986 (Proposition 65) passed by California voters.	Limits discharge of certain chemicals into drinking water.
1988	Amendments to FIFRA, referred to as "FIFRA lite."	Strengthened EPA authority on reregistration, fees, expediting registration, disposal of suspended or canceled pesticides.
	EPA revised its policy of residues in food.	Pesticide residues that present cancer risks of less than one in a million now permitted.

expanding set of test protocols. Among the requirements called for are studies on mammalian toxicology (including lifetime animal feeding studies), environmental chemistry (persistence, mobility, etc.), and effects on fish and other wildlife. New product registrations are granted only after EPA scentists and administrators are satisfied that use of the product does not pose unreasonable risk to humans or hazard to the environment. All aspects of the environment are considered. Soil, air, and water (streams, lakes, oceans, rivers, marshes, and underground aquifiers) are matters of environmental concern, as are the living organisms that reside therein. The protocols employed to ensure safety are complex and not infrequently are at the boundaries of scientific capability. The registration of any new product is a highly complicated and expensive procedure.

As to the second objective, much the same criteria are used in judging whether or not to allow continued use of previously registered pesticides. With the 1972 Amendments to the FIFRA, EPA received authority to forbid or restrict use of pesticides that present unreasonable hazards. The mechanism used to accomplish this end is the RPAR (Rebuttable Presumption Against Registration) procedure, established by EPA in 1975. The RPAR is a stepwise process that begins with EPA's statement of the hazards supporting its contention that the use of the pesticide should be discontinued. These arguments then may be rebutted by any interested parties. Final determination is made by EPA after consideration of information on both risk and benefit. Between 1976 and 1979 certain registrations for a number of chlorinated pesticides (Dieldrin, Aldrin, heptachlor, chlordane) were canceled through the RPAR process. In other cases, restrictions were placed on the use of pesticides to limit exposure to hazards disclosed through the RPAR process.

In 1988 there were amendments to FIFRA requiring EPA to accelerate the *reregistration* of all pesticides registered before 1984. By current standards, earlier test protocols were considered to be insufficient to satisfy the safety concerns of the present. Data gaps, toxicological and environmental, in the earlier registrations were required to be eliminated to permit reregistration of hundreds of currently used compositions.

Residues of pesticides on food commodities also are regulated by EPA although the statutory basis on this control stems from the Federal Food, Drug, and Cosmetic Act (most of which is administered by the Food and Drug Administration). Because the use of pesticides on crops often leads to human consumption of food residues, a complex regulatory mechanism has evolved to determine residue levels that may be presumed safe for such consumption. These levels, termed "tolerances," are defined after consideration of chronic feeding studies in laboratory animals and of analytical data, to determine what levels may follow from approved agricultural practices and a presumed daily intake of the toxicant by humans.

Many agricultural chemicals and their formulations that are manufactured in the

United States are used throughout the world. Although the application of science and technology is at a high level in the United States, and comprehensive regulatory legislation was developed first and most extensively in this country, all the technically developed nations regulate to some degree the manufacture, sale, and use of agricultural chemicals. Germany, Japan, Great Britain, France, and many smaller technically advanced nations have legislation and regulating agencies in this area. This is also true to a lesser degree of most of the smaller nations of the world and for many of the developing nations. In Europe a generalized EEC regulation body is evolving that recognizes certain national differences. Variations in the degree of regulation between Europe and the United States are constantly diminishing.

The World Health Organization (WHO) and the Food and Agricultural Organization (FOA) play an important role in the international regulation of pesticides. A joint undertaking of these two agencies, the Codex Alimentarius Commission, reviews data on pesiticide residues and toxicity and sets tolerance limits on residues that may be considered safe for purposes of international commerce in agricultural commodities.

In the United States, enormous and complex questions of values and priorities—despite numerous studies and reports—remain unresolved. Many chemicals have been banned, or their use has been seriously restricted. The rate of discovery and commercialization of new pesticide products has slowed dramatically. Resolution of these matters relating to ecology, toxicology, and practical need is critical for the agrochemical industry, for agriculture, for the protection of the environment, and in the interests of public health. New strategies for agricultural practice were the subject of a 1989 National Science Foundation Report that especially espoused the *reduction of chemical input*.

Obsolescence of Pesticidal Compounds

Obsolescence of pesticidal chemicals derives from two main causes. The first of these is a consequence of the universal property of living things to adapt by selection to environmental changes, including foreign chemicals. From a practical standpoint, this is best exemplified by the tolerance of certain strains of the housefly to DDT. Certain resistant strains of the housefly require 2000 times the dose of DDT used on susceptible strains to achieve an LD_{50} (the lethal dose required for a 50% kill of a given population). Compounds used to eradicate insect pests in the 1950s and 1960s (e.g., flies in dairy barns, mosquitoes in tropical still waters, etc.) were frustrated by the organisms' resistances. Higher and higher dosages were used to achieve "control." Nor was resistance confined only to insect pests. In time, mites, fungi, weeds, rodents, and so on, all showed evidence of resistance when acaricides, fungicides, weed killers, and rodenticides, respectively, were employed in ever-increased quantities.

The second cause of obsolescence is the development of more useful chemicals. The term "useful" here covers a host of unrelated properties. The new chemical may be cheaper to produce, easier to apply, or less hazardous (to the applicator, consumer, or the environment), or may possess a greater scope or more intense and extensive biological activity; or, in contrast, it may be highly specific in its biological properties. The chemicals listed in the tables that follow were all largely the result of technology that contributed some advantageous property. In certain cases, subsequent research or misuse may have demonstrated that their advantages were overshadowed by the discovery of undesirable consequences.

The case of DDT is an excellent example of the obsolescence of an important pesticidal chemical (Table 29.3). Many compositions of the chlorinated hydrocarbon group were found that were more highly active—either broadly or narrowly effective—than DDT. Indeed this compound stimulated the research and development for toxaphene, the cyclodiene insecticides DDD, dicofol, and so on, and even some of the fungicides and herbicides.

In the 1960s, after the development of sensitive analytical techniques (mostly chroma-

TABLE 29.3 Obsolescence in the Pesticide Industry, DDT and Its History

	Year
First synthesized, Zeidler, Germany.	1874
Laboratory curiosity until discovery of insecticidal properties, Paul Mueller, Switzerland.	1939
Nobel Prize (Medicine), Mueller.	1948
Mass introduction into agriculture and public health areas.	1949 onward
Proliferation of modern analytical techniques. Recognition of insect resistance.	1950–65
Recognition of residue, physiological, and ecological interrelations.	1960–70
State and federal legislation prohibiting or severely restricting use.	1969–70
Virtually eliminated from use in the United States and advanced countries, DDT continues to be used in certain third world and poorer nations.	1970–90

tography or spectroscopic methods) and concern for the proliferation of environmental hazards, DDT was virtually eliminated from agriculture. It is significant that just two decades after the discoverer of the insecticidal properties of DDT received the plaudits of the scientific world through the Nobel Prize, the chemical had been eliminated from use in the United States and its use much reduced in the world at large.

It soon became evident that new compositions, particularly of similar structure to established pesticides (e.g., newer chlorinated hydrocarbons vs. established members of that class or newer organic phosphates or carbamates vs. the established chlolinterase inhibitors), failed to control the organisms against which they had been designed. This is one form of cross resistance. Another, but related, form of cross resistance is exhibited by compounds whose metabolic breakdown is controlled by a specific biochemical pathway. Another form of resistance is manifested by mutations that render the response to the chemical ineffectual, that is, some change in the receptor or the enzyme that normally is attacked. Flies in Danish dairy barns or ticks in Australia exhibit a high degree of cross resistance to a large variety of insecticides, as do vineyards in France and Germany to a variety of fungicides. There are currently areas in the United States where resistance and/or cross resistance to highly effective and relatively new sulfonyl urea and imidazolinone herbicides is exhibited.

THE COMPOUNDS OF THE AGROCHEMICAL INDUSTRY

Complex structures such as the avermectins (a potent miticide, insecticide, and animal parasiticide), the insecticidal products of *Bacillus thuringiensis*, streptomycin (which is employed against certain plant bacterial infections), rotenone (a nonpersistent insecticide), and gibberellic acid (a plant growth regulatory) all have limited but increasingly practical use in agriculture. By far the majority of pesticides are of simpler structure than these, as a consequence in part of the realities of agricultural economics. Structures like those in Table 29.4 containing sequences of optically active sugars or peptides, *cis–trans*-isomeric multiple bonds, numerous asymmetric centers, or multiple annelated rings present formidable obstacles to the synthetic organic chemist and well-nigh insuperable obstacles to engineers who would try to mass-produce them at low to moderate cost.

Because pesticides frequently owe their activity to their "fit" on some biological surface (e.g., an ezyme or a cell membrane), configuration frequently is decisive for biological activity. Differences in activity resulting

TABLE 29.4 Typical Complex Structures Employed in agriculture

Structure	*Name; process; use*
	Avermectins; produced by fermentation; insecticide, miticide, parasiticide (several related structures).
	Thuringiensin; produced by strains of *B. thuringiensis*; insecticidal endotoxin.*
	Streptomycin; Produced by fermentation; bactericide.
	Gibberellic acid; produced by fermentation; plant hormone.
	Rotenone; extracted from derris root; insecticide, piscicide.

Note multiplicity of asymmetric carbon atoms in all five structures.

**Bacillus thuringiensis* endotoxins: These are proteins and mixtures of proteins that vary among the strains of the organism. No single structure is totally descriptive. These endotoxins release a number of toxic factors in the midgut of insects, which have been characterized and synthesized and vary in size between 27 and 140 KD. Other toxins such as Thuringiensin have been isolated.

TABLE 29.5 Structural Isomerism of DDT

Structure	*LC_{50} Drosophila melanogaster (mg/kg wt)*	*% in Technical DDT*
o,p-isomer	145	10–20
p,p-isomer	1	65–85

from some simple isomeric variations are outlined in the following discussion.

Structural isomerism, such as *ortho*, *meta*, or *para* orientation on a benzene ring, can be reasonably controlled by the synthetic process. At times, the insecticide is defined by the process employed in its manufacture. Early technical DDT was an insecticide that contained both isomers and impurities totaling as many as 14 distinct compounds. The technical product meets a standard defined by a minimum setting point at 89°C. The principal impurity is the *ortho–para* structure, which has very little insecticidal potency (see Table 29.5).

The isomers of DDT illustrate the effects of ring substitution on biological activity. Large differences in activity also exist in *cis–trans* isomers, which are given for two organophosphate insecticides in Table 29.6. Optical isomers of pesticidal compounds also exhibit differences in biological activity. This difference is exhibited in the activity of pyrethroids and in the isomers of benzene hexachloride, where the isomer lindane possesses almost all the useful insecticidal activity of the compound.

In recent years, there has been increasing interest in the ecologically "soft" synthetic pyrethroids. These molecules possess chiral carbon atoms, and the configuration at

TABLE 29.6 *Cis–trans* Isomerism and Biological Activity[a]

Structure[b]	*Name*	*LD_{50} Housefly, mg/kg*
	cis-Mevinphos	0.27
	trans-Mevinphos	23
	cis-Diethyl β-chlorovinyl phosphate	1
	trans-Diethyl β-chlorovinyl phosphate	80

[a]Based on Lichtenthaler, F. W., *Chem. Rev.*, **61**, 630 (1961).
[b]In both organophosphate insecticides, the *cis* isomer is about 80 times as active as the *trans* isomer as measured by topical application to the housefly.

these centers largely determines the degree of biological activity. Illustrative of this structural specificity and potency is the fact that for control of important lepidopterous cotton pests only about 0.1 to 0.2 lb/acre of synthetic pyrethroid is required, compared to about 1 to 2 lb/acre of conventional organophosphate or chlorinated hydrocarbon insecticides. Because these pyrethroids have low environmental persistance and are used at such low dosages, the environmental burden is much reduced. Some of the newer herbicides, such as the sulfonyl ureas and imidazolinones, to be discussed later, provide effective weed control in doses of *a few grams* per acre.

It can be concluded from these examples, as well as from the structures that follow, that relatively minor changes in chemical substitution or configuration can cause enormous differences in activity. These structural characteristics do provide some information about the requirements of the receptor site of the biological system affected.

The molecular weight of most agrochemicals provides a simple generalization of the size and complexity of the molecules that are commercially employed as pesticides. The common names of some extensively used pesticides are followed by their molecular weights in parentheses: DDT (354.5), 2,4-D (249), captan (300.6), malathion (330.4), atrazine (215.7), carbaryl (201.2), dicamba (221), glyphosphate—salt (228.2), fenvalerate (419.5), chlorsulfuron (357.8), and assert (288.5). These are small to medium-sized organic molecules, frequently containing halogen, nitrogen, sulfur, and phosphorus, as well as carbon, hydrogen, and oxygen.

Comprehensive descriptions of contemporary pesticides are provided by the references. What follows here are groups of significant pesticides generally typical of a given class or identifying a given principle and commonly used in world agriculture.

Agrochemical Nomenclature

Because no uniform practice exists for the naming of agrochemicals, and because commercial compounds generally possess more than one designation, some clarification of the nomenclature adopted here is in order. In the tables that follow and in the text, the so-called generic or common name has been used most frequently.

Where a material is a specific chemical composition, there is a systematic chemical name generally consistent with the rule of nomenclature in *Chemical Abstracts*. Such a name often is quite cumbersome. The systematic name for the structure of streptomycin given in the text occupies four or five lines of print and, from the standpoint of the chemist, is less informative than the two-dimensional structural formula that is provided.

Agrochemicals developed by private industry usually are initially given an arbitrary designation (generally containing a number) that is employed during the period of exploration and development. This name may prevail during the period when research teams, farm advisors, and experiment stations work with the material. During the course of this period, a proprietary name generally is selected by the company responsible for the development of a potentially successful pesticide. This proprietary name falls within the purview of the U.S. trademark laws. Foreign trademark names that are selected are not always the same as the U.S. name. The product usually is sold under a variety of trademark designations that denote the company of its source. Then the generic or common name may be selected for the commercialized chemical. Most scientific tests that discuss the structure and biological activity, metabolism, toxicology, and so on, of pesticides employ the generic or common name.

Table 29.7 provides some alternate names for pesticides whose structures are given in the text. For a given pesticide, one of the alternate names may be a more common designation than the generic name. It is obvious that even for a relatively simple structure, such as endosulfan, the systematic chemical name is too cumbersome.

Chlorinated Organic Insecticides

The halogenated organics were among the earliest synthetic organics employed as insecticides (see Table 29.8). Nevertheless, largely

TABLE 29.7 Names of Some Pesticides

Generic Name	*Systematic Name*	*Proprietary Name*[a]	*Other Names*
dimethoate	*O,O*-Dimethyl *S*-(*N*-methylcarbamoylmethyl) phosphorodithioate	Cygon[R]	Rogor[R] (Europe) Foslion[R]
carbaryl	1-Naphthyl *N*-methylcarbamate	Sevin[R]	
naled	1,2-Dibromo-2,2-dichloroethyl dimethyl phosphate	Dibrom[R]	RE-4355
captan	*N*-Trichloromethylmercapto-4-cycloyexene-*cis*-1,2-dicarboximide	Orthocide[R]	
endosulfan	6,7,8,9,10,10-Hexachloro -1,5,5a,6,9,9a-hexahydro-6,9-methano-2,4,3-benzodioxathiepin-3-oxide	Thiodan[R]	Malix[R]
glyphosate	Isopropylamine salt of *N*-(phosphonomethyl)glycine	Roundup[R]	
permethrin	3-Phenoxyphenyl methyl (IR,S)-*cis,trans*-3-2,2 dichlorophenyl)2,2-dimethyl cyclopropane carboxylate (approx. 60% *trans* 40% *cis*)	Ambush[R]	Pounce[R]
chlorsulfuron	2-Chloro-*N*(4-methoxy-6-methyl-1,3,5 triazene-2-ylamino carbonyl) benzene sulphonamide	Glean[R]	

[a]Proprietary names generally have the initial letter capitalized and are followed by the symbol [R] as in Sevin[R]; common names are written in lowercase letters.

because of toxicological hazards, both environmental and human, these compounds have been interdicted for sale and manufacture in most of the developed countries. They are still employed, however, in certain countries because of their low cost, and many for their simplicity of manufacture and biological activity. It is important to discuss them here because (1) much of our knowledge of environmental interactions in soil, water, air, and living organisms is derived from our concerns about the metabolism of these compounds; (2) residues of these compounds, even where legal use was forbidden during the last decades, do remain, and hence require our attention; (3) our experience with resistance, particularly insect resistance and cross resistance, was derived from the use of these compounds; and (4) much agrochemical science was derived from the synthethis, analytical methodologies, formulation, and biochemistry of these compounds and even today influences the development of newer active materials.

DDT exhibits a broad spectrum of insecticidal activity, and its success stimulated the research that led to many of the other synthetic organic chlorine compounds. DDT has a relatively moderate acute mammalian toxicity—about 250 mg/kg for the technical product. It is oil-soluble and very insoluble in water—about one part per billion (1 ppb).

In the 1950s and 1960s there was on the one hand increasing evidence of resistance to DDT by various insect species, and on the other hand a proliferation of chemical and analytical techniques enabling quantitative pesticide analyses in complex substrates down to, and in certain cases below, the ppb level. These analyses conclusively showed that DDT and its chief metabolite DDE (dichlorodiphenyldichloroethylene) persisted in the environment, principally in cold and certain temperate climates, and concentrated in fatty tissue, chiefly in organisms at the top of any given ecosystem.

There is considerable, but variable, cross resistance from DDT to other chlorinated hydrocarbons such as benzene hexachloride and the cyclodiene insecticides. However, when one substitutes a *para*-methoxy group for *para*-chlorine in the DDT molecule, a far less persistent—and somewhat less active—but

TABLE 29.8 Group 1—Miscellaneous Chlorinated Hydrocarbons

Name	*Structure*	*Significant Intermediates*	*Uses—Properties*
DDT		Chlorobenzene, chloral.	General insecticide (malaria control).
Toxaphene	Complex mixture	Chlorinated camphene (68–69% chlorine).	Insecticide (cotton).
BHC Lindane (99.5% pure γ-isomer)		Benzene, chlorine ($h\nu$).	Insecticide used heavily in Japan (rice).
Hexachlorobenzene		Benzene, chlorine.	Fungicide (cereal-seed treater).
Dicofol		DDT, chlorine (hydrolysis)	Acaricide; miticide.

useful insecticide, methoxychlor, is obtained. The substitution of OH for H in the trichloroethane group greatly reduces and almost eliminates the general insecticidal properties of the DDT configuration. Nevertheless, such molecules possess high acaricidal and ovicidal properties (e.g. dicofol).

DDT metabolite, DDE

The simplicity of the chlorination of benzene led to the discovery of the usefulness of benzene hexachloride and lindane as insecticides—and of hexachlorobenzene, the true hexachlorobenzene, as a grain fungicide. (Benzene hexachloride is a misnomer for a chemical. Benzene hexachloride, originally prepared by Michael Faraday in 1825, is hexachlorocyclohexane.) Lindane has the spatial configuration that possesses the (99% + gamma isomer) highest activity: six hexachlorocyclohexanes are found in commercial preparations although the stereochemistry of the cyclohexane ring does provide many more theoretical isomers. Lindane has three axial and three equatorial chlorines.

The Cyclodiene Insecticides

Hexachlorocyclopentadiene (C_5Cl_6) itself possesses herbicidal activity. As a diene combined with a variety of dienophiles and modified thereafter, it has given rise to a large number of very potent insecticides.

Hexachlorocyclopentadiene

The earliest of these was chlordane, which was employed both in agriculture and for the control of home and garden pests. Chlordane is somewhat herbicidal as well, and found some use for crabgrass control.

Aldrin, dieldrin, heptachlor, and endrin are all very highly active toxicants with moderate to highly acute mammalian toxicity, with a wide spectrum of insecticidal activity. They and their metabolites tend to concentrate in fatty tissue and have considerable persistence in the environment. Federal regulations (RPAR process; see preceding section on government regulation) as well as state legislation have either eliminated or severely limited their agricultural application. The names, uses, and structural formulae of several of these chemicals are given in Table 29.9.

Despite the fact that this group of chemicals has had wide agricultural and other usage, little has been unequivocally established concerning the essential biochemistry involved in their activity. Recent biochemical studies reveal sodium channel and GABA interactions for insect and mammalian organisms for the cyclodienes.

Whereas one of the mechanisms of resistance to DDT can be at least partially related to the following enzymatic reaction:

$$\text{DDT} \xrightarrow[\text{dehydrochlorinase enzyme}]{-\text{HCl}} \text{DDE}$$

no such simple mechanism can be associated with cyclodiene resistance mechanisms. Frequently, in the field, some insect species exhibit a generalized chlorinated hydrocarbon resistance.

Derivatives of Organic Aids

As many biological systems involve the use of natural acids, it is not entirely unexpected that organic acids or their derivatives find use in the agrochemical industry (Table 29.9). For many years the dominant organic acid herbicides were 2,4-D (2,4-dichlorophenoxyacetic acid) and 2,4,5-T (2,4,5-trichlorophenoxyacetic acid). In their mode of action they resemble the natural plant auxin, indoleacetic acid, and are used primarily to control broad leaf weeds. The worldwide use of 2,4,5-T has been severely restricted primarily because of a trace quantity of an impurity, 2,3,7,8-tetrachlorobenzodioxin (below).

2,3,7,8-Tetrachlorobenzodioxin

This substance is very toxic to mammals, and it also possesses carcinogenic, mutagenic, and teratogenic effects. The use of 2,4-D and its derivatives also has been restricted because of drift of the herbicide to nontreated areas resulting from its appreciable volatility at application temperatures.

Many other aryloxy acid derivatives have been commercialized as herbicides, however, including types that do not act as auxin mimics in the plant. One of these is napropamide, which inhibits root growth primarily in grassy weeds. Other derivatives are a group of important phenoxyphenoxypropionic acids, first developed by Hoechst AG. This group includes types such as diclofop-methyl and haloxyfop-methyl, which are as inhibitors of fatty acid biosynthesis and find use as postemergence herbicides for control of annual grassy weeds. Important herbicides such as alloxydim and sethoxydim, have a similar mode of action to this even though their structures are quite different.

A wide variety of aromatic acids have been used as herbicides. Some of these, such as amiben, dicamba, and picloram, mimic auxin in their mode of action, similarly to 2,4-D; however, they possess chemical, biochemical, and physical characteristics that make them more suitable for particular applications than 2,4-D is. Expiration of patent protection for many of the aromatic acid herbicides has led to their manufacture by a number of companies.

TABLE 29.9 Agrochemicals Derived from Organic Acids (continued)

Name	*Structure*	*Major Intermediates*	*Uses*
2,4-D		2,4-Dichlorophenol, chloroacetic acid	Broadleaf herbicide
2,4,5-T		2,4,5-Trichlorophenol, chloroacetic acid	Broadleaf herbicide
Napropamide		Naphthol, chloropropionic acid, diethyl amine	Preemergence herbicide
Diclofop-methyl		Phenoxyphenol, chloropropionic acid	Postemergence herbicide
Haloxyfop-methyl		Pyridyloxyphenol, chloropropionic acid	Postemergence herbicide
Chloramben		Reduction of the dichloronitrobenzoic acid	Preemergence herbicide
Dicamba		Methylation of the dichlorosalicylic acid	Postemergence herbicide
Picloram		Chlorination of 2-picoline, ammonia, hydrolysis	Postemergence herbicide
1-Naphthalene acetic acid		Naphthalene, chloroacetic acid	Plant growth regulator
1-Naphthalene-acetamide		The above acid and ammonia	Plant growth regulator

TABLE 29.9 Agrochemicals Derived from Organic Acids (continued)

Name	*Structure*	*Major Intermediates*	*Uses*
Acifluorfen		Dichlorobenzotrifluoride, the hydroxynitrobenzoate ester	Postemergence herbicide
Imazaquin		2,3-Quinolinedicarboxylic anhydride, 2-methyl-leucine amide and ring closure	Soybean-selective herbicide
Alachlor		2,6-Diethylaniline, formaldehyde, and chloroacetyl chloride	Corn-selective preemergence herbicide
Metolachlor		2-Ethyl-6-methylaniline, 2-bromo-1-methoxypropane, chloroacetyl chloride	Corn-selective preemercence herbicide
Propanil		3,4-Dichloroaniline, propanoyl chloride	Weed control in rice
Deet		*m*-Toluoyl chloride, diethyl amine	Insect repellent
Metalaxyl		2,6-Diethylaniline, chloropropionic acid, methoxyacetic acid	Systemic fungicide
Diflubenzuron		2,6-Difluorobenzamide, *p*-chloroaniline, phosgene	Insecticide, chitin inhibitor
Iprodione		3,5-Dichloroaniline, glycine phosgene, isopropyl amine	Systemic fungicide

TABLE 29.9 Agrochemicals Derived from Organic Acids (continued)

Name	*Structure*	*Major Intermediates*	*Uses*
Captafol		Maleic anhydride, butadiene ammonia, the sulfenyl chlorine	Fungicide
Captan		Maleic anhydride, butadiene, ammonia, carbon disulfide, chlorine	Protectant fungicide
Chlorphthalim		Tetrahydrophthalic anhydride, *p*-chloroaniline	Turf herbicide
Oxadiazon		The hydrazine, trimethylacetyl chloride, phosgene	Preemergence and early postemergence herbicide

The increased competition for markets for these materials has increased their use and reduced their cost to the user.

Another group of compounds, which are exemplified by 1-naphthaleneacetic acid and 1-naphthaleneacetamide, are auxin mimics and find use as plant growth regulators. They stimulate root formation in cuttings and transplants.

As we have seen, herbicides can cause plant death in a variety of ways. Another important aromatic acid herbicide is aciflorfen, one of a group of peroxidizing diphenyl ether herbicides. These types promote the generation of free radicals so that important plant systems are injured. Aciflorfen is an important herbicide used in soybeans and rice.

Imazaquin, an example of the imidazolinone class of herbicides, is particularly selective to soybeans. These compounds have a similar mode of action to that of the sulfonylurea herbicides, which will be discussed later. They block branched chain amino acid biosynthesis by inhibition of the enzyme acetolactate synthase (ALS).

The chloroacetanilide herbicides such as alachlor and metolachlor are useful as preemergence herbicides in the cultivation of corn. They are characterized as growth inhibitors by preventing cell enlargement. Chloroacetanilide type compounds are well known as alkylating agents; so it is not surprising to find that alachlor is a carcinogen and that its use has been greatly restricted.

Another group of anilide herbicides is exemplified by propanil, whose mode of action is quite different from that of the chloroacetanilides. Members of this large class of compounds, which also includes various ureas and carbamates, affect the photosynthesis apparatus of the plant. They will be discussed in more detail later.

Organic acids and their derivatives find use in many other applications outside the herbicide area. Deet, *N*,*N*-diethyl-*m*-toluamide, is widely used as an insect repellent. Another toluamide is meopronil, a fungicide that controls rice sheath blight. The natural and synthetic pyrethroid insecticides are acid esters, which are discussed more fully later.

Metalaxyl is a member of a class of compounds that structurally resemble the chloroacetanilides; however, it is a systemic fungicide. Diflubenzuron is representative of a class of insecticides that owe their action to inhibition of chitin synthesis by the developing insect. Metalaxel and diflubenzuron are good examples of compounds prepared by the synthesis chemist as herbicide candidates that have commercial levels of activity in entirely different applications. Such examples point out the need for a balanced synthesis and screening program for the development of new agrochemicals.

Cyclic imide types include herbicides, insecticides, and fungicides. Iprodione is an example of the fungicidal imides. These systemic fungicides have a 3,5-dichlorophenyl group attached to the imide nitrogen. Captan is an example of a protectant imide fungicide. Compounds related to captan have a haloalkylthio group attached to the imide nitrogen. The haloalkyl group attacks several different biochemical sites in the fungi; so major resistance has not deveoped to this class of fungicides. Unfortunately, similar bochemical reactions occur in nontarget species and lead to mutagenic complications in test animals.

Imide herbicides are exemplified by types similar to chlorphthalim and oxadiazon, which are expected to have a mode of action similar to acifluorfen. A cyclic amide herbicide is clomosone, a bleaching herbicide. These types will be discussed more fully later. Phosmet, an example of an insecticidal imide, also is an organophosphate and will be further discussed with the phosphorus compounds.

Organic Phosphorus Agrochemicals

Historically organic phosphorus agrochemicals meant insecticides. Today that is not the case. Organic phosphorus agrochemicals (Table 29.10) now include most applications. Just as organic carboxylic acids are vital to various life processes, so too are phosphorus compounds. Thus it is not surprising that phosphorous compounds find their place as versatile agrochemicals.

Most of the phosphate insecticides owe their biological action to their inhibition of the cholinesterase enzymes of the insect. This inhibition causes a buildup of acetylcholine in nerve tissue, resulting in uncontrolled energy trasmission through the nerve, which is manifest as convulsive activity at vital muscular centers.

Acetylchlorine chloride

Because all animal life with a nervous system uses this nerve impulse transmission system with acetylcholine, enzyme inhibitors that find their way to the nerve sites have their effect. For that reason many of the first-generation phosphate insecticides, such as parathion, had a high mamalian toxicity in addition to their insecticidal action. In response to the need to develop safer products, mamalian toxicity became just as important as insecticidal action in the development of new insecticides. Compounds such as malathion have an ester function that can be saponified by enzymes in mammals before it reaches vital systems. Insects are not able to facilitate this detoxification as well as mammals, and in this case differences in rates of detoxification between insects and mammals led to a safer compound.

Environmental degradation is a useful means of reducing the level of an agrochemical in an agricultural product, and it can be monitored by residue studies that determine the level of agrochemical contamination in food products. For compounds that quickly degrade in the environment, fruit and vegetables can be harvested only a few days after spraying of the agrochemical. Compounds that degrade more slowly necessitate longer periods before harvesting of the food product; so compounds that degrade only slowly in the environment are restricted to those applications that are not likely to cause contamination problems. Phosmet is an example of an insecticide that finds major use in control of the coddling moth in apples. The insecticide is applied to the apple trees when they are in bloom, when the moth lays the eggs in the

TABLE 29.10 Important Organophosphates

Name	*Structure*	*Major Intermediates*	*Uses*
Parathion		*p*-Nitrophenol, phosphorus trichloride	Insecticide
Bolstar		*p*-Methylthio-phenol, propyl mercaptan, phosphorus trichloride	Foliar insecticide
Malathion		Ethyl maleate, dimethyl dithiophosphoric acid	General insecticide
Dimethoate		Chloroacetic acid, methyl amine, dimethyl dithiophosphoric acid	Systemic insecticide
Diazinon		The hydroxypyrimidine diethylthiophosphoryl chloride	General insecticide
Glyphosate		Glycine, formaldehyde, phosphorus acid, isopropyl amine	Nonselective postemergence herbicide
Orthene		Dimethyl thionophosphorochloridate, ammonia, acetic anhydride	Systemic insecticide
Fosetyl-Al		Ethyl phosphite, aluminium salt	Systemic fungicide
Ethrel		Phosphorus trichloride, ethylene oxide	Plant growth regulator
Phosmet		Phthalimide, formaldehyde, dimethyl dithiophosphoric acid	Insecticide for a variety of crops
O,O-Diethyl-*O*-phenyl-phosphorothioate		Diethylthiophosphoryl chloride, phenol	"Extender" for thiocarbamate herbicides

flowers. At harvest, much later, the apples are free of worms.

During the early development of insecticides, compounds were discovered that had the interesting property of movement in the plant from the point where they were applied. This systemic character is very useful in the control of certain insects and mites that are found in hard-to-spray areas of the plant. Dimethoate, which is an example of such a systemic material, also has a high degree of selective toxicity. However, in addition to being toxic to many insects and mites it also is fairly toxic to some birds. The cholinesterase inhibitor is not the dimethoate itself but rather is its oxidized phosphoro metabolite. Many mamalian systems hydrolyze the amide group before activation of the phosphorus. Then the acid is rapidly excreted from the body.

This systemic character is also one of the prime characteristics of another phosphorus containing agrochemical, glyphosate, which is a translocated postemergent herbicide. The herbicide is applied to the leaves and is translocated down into the roots of the plant. Plants with extensive root systems often return after application of a standard postemergent material. Glyphosate moves down into the roots through the phloem system of the plant, where it inhibits the plant's synthesis of aromatic amino acids so that the plant gradually dies. If rapid weed control is required, other herbicides such as paraquat may be required.

Another interesting systemic phosphorus compound is fosetyl Al, a systemic fungicide. It has preventive and curative activity against many Oomycetes on many different crops. This combination of a phosphite and aluminium gives the necessary chemical and physical properties to the material to make it an effective fungicide.

Just as with the chlorinated hydrocarbons, resistance has developed by many types of insects to the organophosphate insecticides. In retrospect this is not altogether unexpected. The phosphate insecticides' mode of action involves inhibition of the cholinesterase enzymes of the insect. When the level of insecticide spraying is such that nearly all of the insects are killed in the applied area, those insects that survive and reproduce with other surviving insects then require even higher levels of insecticide for their control. With repeated applications of similar-acting materials, resistance gradually develops to these materials. For those insects and mites that have a relatively short life cycle, several generations can occur in one year. It is particularly with these organisms that insecticide resistance has become a major problem. There now are mosquitoes, aphids, and mites in some areas that are resistant to most insecticides. Development of resistance is not limited just to insecticides. Any material that owes its biological effect to activity on one type of active site can expect resistance to develop. This applies to insecticides, miticides, bactericides, fungicides, and herbicides. Any agrochemical application plan should then take resistance development into account in its use. The rotation of agrochemicals now is a commonly accepted strategy to ameliorate resistance to specific agrochemicals.

Carbamate Agrochemicals

Carbamates are not common natural biological constituents, but carbamates of various types find use as agrochemicals. Biological systems do, however, widely use amide and ester materials. Both amide and ester functions are found in the carbamate molecule.

Generalized carbamate structure

Carbamates (Table 29.11) in the common insecticide types owe their mode of action to cholinesterase inhibition. Similar to the organophosphate insecticides, they react with the cholinesterase enzyme much as the phosphates do; instead of phosphorylation of the enzyme, carbamoylation occurs. In this way their modes of action are very similar. In the insecticide types R_1 can be a wide variety of different groups. One of either R_2 or R_3 is

TABLE 29.11 Carbamate Agrochemicals

Name	*Structure*	*Major Intermediates*	*Uses*
Carbaryl		Naphthol, methyl isocyanate	Insecticide
Methomyl		Acetyl chloride, methyl mercaptan, hydroxyl amine, methyl isocyanate	Contact insecticide
Maneb		Ethylene diamine, carbon disulfide, manganese salt	Fungicide
EPTC		Dipropyl amine, phosgene, ethyl mercaptan	Preemergence herbicide
CDEC		Diethyl amine, carbon disulfide, 2,3-dicyloropropene	Preemergence herbicide
Thiodicarb		Methyl mercaptan, acetonitrile, hydroxyl amine, methyl isocyanate, and sulfur dichloride	Insecticide
Propham		Aniline, phosgene, isopropyl alcohol	Pre- and postemergence herbicide
Chlorpropham		*m*-Chloroaniline, phosgene, isopropyl alcohol	Pre- and postemergence herbicide
Metam-sodium		Methyl amine, carbon disulfide, caustic	Soil sterilant, fungicide, nematocide
Benomyl		Phenylenediamine, phosgene, butyl amine, methanol	Systemic fungicide

usually hydrogen, and the other is usually methyl. In some cases various groups replace the hydrogen to give compounds in which the physical and chemical properties are modified to give more suitable characteristics, such as selectivity, mamalian toxicity, and increased insecticidal activity. These groups are removed by hydrolysis before the carbamate expresses its action on the enzyme system. This method of delivery of the active compound is called the "pro-drug approach." Thiodicarb is an example of such a compound. Its major use is in cotton and soybeans.

Carbamates also include many compounds that have herbicidal activity. Important carbamate herbicides are propham and chloropropham. They are not effective cholinesterase inhibitors, but they have been shown to inhibit several other important systems such as RNA synthesis, protein synthesis, mitosis, oxidative phosphorylation, and the Hill reaction of photosynthesis.

The oxygen atoms of the carbamate can be replaced by sulfur atoms to give important agrochemicals.

Dithiocarbamate types

Historically compounds such as CDEC were important to show that herbicidal activity could be found in compounds where the R groups were alkyl. Unfortunately, toxicology problems were found with the better herbicidal compounds.

The metallic salts of ethylene-bis-dithiocarbamate have proved to be important foliar fungicides, serving as broad spectrum foliar fungicides. These compounds are protectant materials with little systemic action. Also they show their action at multiple sites, and for that reason resistance is less prevalent for these materials. Unfortunately, compounds that attack multiple sites also tend to have possible toxicological problems.

The monothiolcarbamates have been shown to have wide application as herbicides, several having been developed for a variety of applications as preemergent herbicides.

Monothiolcarbamate types

The discovery that crop tolerance for a particular herbicide could be enhanced has led to the identification of many compounds that "antidote" the herbicidal effect of the thiolcarbamate herbicides. *N,N*-Diallyldichloroacetamide was the first material commercialized with thiocarbamate herbicides for use in corn. This same concept also has application for other types of herbicides. Also it was discovered that repeated application of the same thiolcarbamate led to rapid microbial degradation of the material in the soil. Addition of "extenders" to the thiolcarbamate herbicide counteracted with degrading effect of the microorganisms in the treated soil.

Urea Type Agrochemicals

Substituted urea compounds are another group of agrochemicals that, like the carbamates, are not common constituents of biological systems. However, the carbonyl nitrogen structural fragment is widely found in biological systems. Most of the commercial urea agrochemicals are herbicides although, as might be expected, other types of biologically active materials are also found in this class.

The first generation of urea herbicides is represented by types such as monuron and linuron (see Table 29.12). These compounds in general interfere with the photosynthetic process, as they inhibit electron transport in the photosystem II process of the plant. They have varying degrees of selectivity, depending upon the structure. Some owe their selectivity to physical properties such as water solubility and soil absorption. If the crop is a deep-rooted perennial, the shallow rooted annual weeds are controlled by using a herbicide that is not easily leached down into the root zone

TABLE 29.12 Urea Herbicides

Name	*Structure*	*Major Intermediates*	*Uses*
Monuron		*p*-Chloroaniline, phosgene, dimethyl amine	Preemergence herbicide
Linuron		3,4-Dichloroaniline, phosgene, hydroxylamine, dimethyl sulfate	Herbicide
Chlorsulfuron		2-Chlorobenzenesulfonyl isocyanate, the aminotriazine	Herbicide in cereals
Chlorimuron ethyl		The sulfonamide, phosgene, the aminopyrimidine	Soybean herbicide
Tebuthiuron		The aminothiadiazole, methyl isocyanate	Herbicide

of the crop. Other urea herbicides such as linuron owe their selectivity to degradation of the herbicide by the crop plant.

A subclass of the urea herbicides includes the sulfonylurea herbicides pioneered by Du Pont, which are among the most active herbicides discovered. Their activity against certain plants is at a rate of a few grams per acre. Chlorsulfuron (see Table 29.12) is an example of a sulfonylurea herbicide. The breakdown of chlorsulfuron in the soil is greatly affected by pH, soil organisms, rainfall, and temperature. Degradation of certain sulfonylureas can occur in a period of a few hours under the proper conditions. In other soils, degradation can take place very slowly so that there may be carryover into the next growing season, and injury may occur to sensitive crops. This class of herbicide owes its mode of action to inhibition of the enzyme acetolactate synthase (ALS) a key enzyme in the plant's synthesis of branched chain amino acids. Sulfonylurea crop selectivity appears to be due solely to rapid metabolic breakdown of the herbicide by the tolerant crop. Sensitive plant species metabolize the herbicide much more slowly than tolerant plants.

A key exception to the rule that sulfonylurea herbicidal selectivity is due to metabolic inactivation is the finding that plants through mutation, selection, or genetic engineering methods have acquired a herbicide-resistant form of the ALS enzyme. Crop species that have been selected or engineered for resistance to sulfonylureas includes soybeans and tobacco. Similarly, certain weeds that have developed by natural mutation and selection have a high degree of selectivity to this class of herbicides. This again points out the common problem of development of resistance to those materials that have a single mode of action for their biological effect.

Heterocyclic Agrochemicals

This is a broad grouping of many different types of agrochemicals (see Table 29.13). The classes of agrochemicals previously discussed often include heterocyclic types. There are heterocyclic acid types such as picloram, heterocyclic phosphates such as diazinon, heterocyclic carbamates such as benomyl, and heterocyclic ureas such as tebuthiuron.

Triazine compounds such as atrazine have found major use in weed control in corn. Corn possesses an enzyme that degrades triazines such as atrazine more effectively than the weeds that compete with it.

Quarternary pyridyl compounds such as diquat and paraquat are important post-emergence herbicides. They are fast-acting nonselective herbicides, which owe their activity to their ability to produce a free radical inside the chloroplast of the plant that intercepts the transfer of energy in the early stages of the photosynthetic process (photosynthesis system I). This free radical is not the actual toxic agent but generates the peroxide type agents formed by the free radical. These herbicides are particularly useful because they are bound in the soil, particularly by certain types of clays, into a form that is not available to plants. Seeds planted in this treated soil are not injured by the bound herbicide.

Systemic fungicides are a fertile area for heterocyclic compounds. Fungicides such as benomyl and carboxin were early systemic compounds. An important property of these systemic materials is their ability to move through the plant and control the fungal organisms in places that cannot be reached by the typical preventive nonsystemic fungicide. However, these systemic materials have a single mode of action, and repeated applications lead to resistance development by the organisms.

Triazole and imidazole derivatives have been a popular source of systemic fungicides. These compounds, such as triadimefon and prochloraz, inhibit various steps in the fungal organism's synthesis of ergosterol. These compounds are also single-mode-of-action materials, and, as might be expected, resistance is building to these useful compounds. This again points to the need of a use strategy that minimizes the potential for buildup of resistance.

"Bleaching" herbicides is another type that is heavily represented by heterocyclic compounds. Norflurazon and fluridone are examples of these heterocyclic herbicides. these compounds exhibit their action by interrupting the biosynthesis of carotene in the plant. This leads to the destruction of the chlorophyll in the plants, which turn a characteristic white or pink color.

Synthetic Pyrethroid Insecticides

The natural insecticide, pyrethrum traditionally has been considered to have the qualities of the ideal insecticide because it is a powerful insecticide and harmless to mammals under normal use conditions. However, the natural pyrethroids are too unstable outdoors to find use as agricultural insectides. In 1949 the first synthetic pyrethroid, allethrin, was announced (see Table 29.14). It acted more rapidly and effectively than the natural material, particularly on the common housefly; but on other species it was less active. This initiated the search for related materials with a combination of properties that would eventually lead to commercial insecticides that could be used in outdoor agricultural applications. The discovery of permethrin by Michael Elliot in 1973 showed that pyrethroids could be found that were stable to outdoor environments. Permethrin also had exceptionally high insecticidal action (0.1 lb ai/A).

About this same period it was becoming increasingly evident that insect resistance was building to the chlorinated hydrocarbon, phosphate, and carbamate insecticides. The new environmentally stable synthetic pyrethroids such as permethrin and fenvalerate found ready acceptance in cotton, fruit, and vegetable crops.

A new generation of even more potent synthetic pyrethroids is now upon us, and these new products are at least ten times more active than permethrin. These new materials require only 0.01 lb ai/A for insect control,

TABLE 29.13 Heterocyclic Agrochemicals

Name	*Structure*	*Major Intermediates*	*Uses*
Atrazine		Cyanuric chloride, ethyl amine, isopropyl amine	Herbicide
Paraquat		Pyridine, methylation agent	Nonselective postemergence herbicide
Bromacil		Acetoacetic ester, *sec*-butyl amine, phosgene, bromine	Herbicide
Bentazon		Anthranilic acid, isopropyl amine, sulfuryl chloride	Postemergence herbicide
Carboxin		Aniline, ketene, mercaptoethanol	Systemic fungicide
Triadimefon		Bromopinacolone, *p*-chlorophenol, bromine, 1,2,4-triazole	Systemic fungicide
Prochloraz		Trichlorophenol, hydroxyethyl propyl amine, phosgene, imidazole	Systemic fungicide
Clomozone		2-Chlorobenzaldehyde, hydroxylamine, chlorotrimethylacetyl chloride	Herbicide
Norflurazon		*m*-Trifluoromethylphenyl hydrazine, mucochloric acid, methyl amine	Herbicide
Fluridone		*m*-Trifluoromethyl benzaldehyde, acetophenone, methyl amine	Herbicide

TABLE 29.14 Synthetic Pyrethroid Insecticides

Name	*Structure*	*Uses*
Allethrin		Knockdown agent for flies
Pemethrin		General insecticide
Fenvalerate		General insecticide
Cyfluthrin		Insecticide
Fenpropathrin		Acaricide
Deltamethrin		Very potent insecticide
Flumethrin		Control of cattle ticks
Bifenthrin		Insecticide, miticide, aphicide
Tefluthrin		Insecticide against soil insects

TABLE 29.14 Synthetic Pyrethroid Insecticides (continued)

Name	*Structure*	*Uses*
Fluvalinate		Insecticide, control of spider mites; nontoxic to honeybees
Etofenprox		Insecticide for rice

which is truly remarkable when compared to the 1 to 2 lb ai/A necessary for the phosphate and carbamate insecticides. However, these new materials are much more structurally complex and require multistep synthetic schemes for their preparation. These new compounds are wholly or partially resolved enantiomers and diastereomers.

There are two general classes of synthetic pyrethroids, the halovinylcyclopropane types and the open chain types. Permethrin was the first of the halovinylcyclopropane type and fenvalerate was the first of the open chain type that had outstanding properties.

For the compounds like permethrin a 4-fluoro substituent on the benzyl portion of the molecule, as in cyfluthrin, generally enhances insecticidal potency. Replacing the chlorine in the *trans* acid side chain with a 4-chlorophenyl group, as in flumethrin, gives specific activity against cattle ticks. Bifenthrin is active against insects, aphids, and mites. Telfuthrin has a higher vapor pressure than most synthetic pyrethroids and is active in soil against the corn rootworm. Fenpropathrin was commercialized primarily for its acaricidal action.

The first of the open chain types that was commercialized was the racemate fenvalerate. In 1986 the single (S,S) isomer was commercialized as esfenvalerate. Fluvalinate greatly extends the range of structural variations and biological properties of these active materials. It controls spider mites and is nontoxic to honeybees. Etofenprox is no longer an ester, and it shows promise against rice pests with reduced toxicity to fish, which is a major consideration in Japan for the rice market.

Insect Hormone Compounds

The understanding of the *molecular basis* for the remarkable changes (metamorphosis) of the insect from egg to larvae, to pupa, to adult (with other less dramatic intermediate stages as well) has been a triumph of academic entomological science. It has led to the isolation, characterization, and synthesis of the hormone that induces the molting and the hormone that tends to maintain the insect at any given stage. Researchers know now that there is greater complexity involved in the metamorphic process; that certain peptides regulate the amount of these hormones at each stage, and that, for example, the structure for J.H. given below is only one of a group of similar compounds that different families of insects employ in their specific metamorphoses.

Methyl *trans*, *trans*, *cis*-10-epoxy-7-ethyl-3,11-dimethyl-2,6-tridecadienoate (J.H.)

Ecdysone—moulting hormone

To date no adverse mammalian effects have been observed for J.H. and ecdysone. It is an attractive strategy for industry to provide compounds that mimic these hormones (see Table 29.15). By interfering with the normal equilibrium, one can prevent insects from reaching the sexually mature adult stage or from producing fertile mature eggs that will develop into larvae. This is a strategy then of the hormone approach to insect control. A hormone differs from a conventional insecticide in that its aim is not to kill the insect directly but to interfere with its life cycle.

Interesting applications of juvenile hormone mimics such as methoprene have been developed. These include mosquito control, particularly where the use of toxic pesticides must be avoided; "feed-through" formulations for horn fly control in ruminants, where the flies breeding in the manure fail to metamorphose into sexually mature adults; silk worm larval treatments, where greater-sized cocoons (more silk) are produced; the protection of stored tobacco from a variety of insect pests; the control of certain fly species in mushroom culture; and the control of insects associated with dogs and cats.

A strategy that decreases and in certain cases eliminates the use of mammalian toxics from home application is the strategy of cockroach population control with J.H. mimics. By preventing the cockroach from producing viable eggs, a reduction and in some cases an elimination of cockroach populations can be achieved through the employment of certain of these hormone mimics.

One of the weakness in the J.H. or J.H. mimic approach to field insect control relates to the fact that the larval form is frequently the most voracious and destructive of the insect life stages. Chemicals that would induce premature pupation conceivably would be

TABLE 29.15 Synthetic J.H. and Ecdysone Agonists and Antagonists

Hormone	*Structure*	*Use*
Methoprene		See text
Precocene II		Laboratory and developmental use
Hydroprene		Cockroach population control, homoptera and lepidoptera activity
Fenoxycarb		Lepidoptera, ant, cockroach control
RH 5849		Experimental compound, ecdysone agonist

quite useful, and considerable effort has been expended toward the search for such substances. Indeed, some compositions are known, for example, precocene II, that induce premature metamorphosis, but as yet no useful agricultural application has been developed from such research. Investigation also is proceeding in relation to the hormonal control of other insect vital processes in the search for new vulnerable targets. Nevertheless, a more easily synthesized, smaller molecule within the confines of agricultural economics that possessed one or all of the above physiological properties would indeed provide a breakthrough in insect control.

In Table 29.15 we provide structures of J.H. agonists and antagonists and of a substance with some ecdysone-like agonistic properties.

Dinitrophenols

One of the earliest groups of synthetic organic chemicals, predating World War II and still in use, is the alkyldinitrophenols. The chemicals have broad biological activity embracing the insecticide, fungicide, acaricide, and herbicide fields. This broad activity results from their capacity to block oxidative phosphorylation, thus interfering with vital biochemical synthesis cycles found in both animal and plant life. As a result the EPA has reviewed and is continuing to review the registration of these toxicants. The trend is toward their interdiction except for noncrop and exceptional applications. There are many other nitrophenols and their derivatives with biological activity. Salts, esters, and free phenols all are utilized in commercial preparations. Table 29.16 illustrates dinitrophenol type agrochemicals.

Soil Fumigants

In many areas, the population of nematodes and pathogenic fungi is so high that seeds are destroyed, or the emerging seedling and plant are similarly attacked so that little or no crop can be harvested. Such soil is treated with a fumigant, usually a volatile general toxicant which frequently is injected into the soil prior to planting. These substances are generally, but not exclusively, low molecular weight halogenated compounds. In addition to those chemicals given in Table 29.15, chloropicrin (CCl_3NO_2), methyl bromide (CH_3Br), and other halogenated C_1, C_2, and C_3 hydrocarbons find extensive application in soil fumigants.

In recent years, the toxicological properties of some of these halogenated fumigants have been intensively reinvestigated. DBCP, for example, has been shown to cause mammalian male sterility, and ethylene dibromide is suspected of mutagenic and carcinogenic potential. As a result, their agricultural application is either totally interdicted or very severely limited. A current need exists for *effective, safe* nematicides and soil pathogenic organism sterilants.

Fumigants of various structures also are employed to preserve harvested crops, corn, small grains, and so on, from insects and other pests. Vapam, the sodium salt of *N*-methyl dithiocarbamic acid, decomposes or metabolizes in the soil to methyl isothiocyanate and provides a local fumigation of fungi, nematodes, and certain insects.

TABLE 29.16 Dinitrophenol Type Agrochemicals

Generic Formula	R_1	R_2	*Functions*	*Name*
	$-CH_3$	H	Insecticide Herbicide	DNOC
		H	Insecticide Herbicide	DNOCHP
			Acaricide	Dinocap
	$-\overset{CH_3}{\overset{\vert}{C}}-(CH_2)_5CH_5$	$-\overset{O}{\overset{\Vert}{C}}-CH{=}CH{\cdot}CH_3$	Fungicide	

There are environmental concerns, particularly regarding atmospheric and water contamination associated with these volatiles, that restrict their current usage.

Synergists

It has long been known in medicine and in agriculture that the use of two chemicals can result in additive, less-than-additive (antagonistic), or greater-than-additive biological effects. Where a greater-than-additive result occurs, the two chemicals are said to behave synergistically. In the pesticide industry, there are a number of moieties that have little or no activity themselves but that when used with a toxicant improve the activity of the toxicant. These synergists are most advantageously employed when the toxicant is an expensive chemical, for such a combination possesses a decisive, practical economic advantage.

The methylenedioxyphenyl synergists have the following general structure:

Here R_1, R_2, and R_3 have a multiplicity of organic functions, and these materials synergize many insecticides. In piperonyl butoxide, an important pyrethroid synergist, the R_1 group is butoxyethoxyethyl. In addition there also is a propyl group in the 6-position in piperonyl butoxide. Certain alkyne, heterocycyle, imide, and thiocyanato linkages also possess degrees of synergy with insecticides. Synergists find considerable use in aerosol sprays with pyrethrins or synthetic pyrethroids such as allethrin.

Patents frequently are written for the concurrent use of two or more toxicants where greater activity than expected results from the combination. Thus, for example, the employment of a mixture of two herbicides may provide a synergistic increase of activity.

Pheromones and Semiochemicals

The current pressure to reduce the quantity of toxicants in the environment has stimulated greater interest in the use of pheromones.

Many chemicals responsible for the sexual stimulation of certain male or female insects have been isolated and characterized. These are usually fairly long chain hydrocarbons, frequently olefinic and often with carbonyl or alcohol or ester terminal functions (see Table 29.17). In addition to sex attractants, there are other types of signal agents. There are chemicals that simulate attractive food entities, danger, and various social insect functions. As with the hormone approach, there appears to be less ecological hazard with the use of pheromones that with more traditional insecticides. Some success has been achieved, usually in locations where the migration of insects is difficult. As an example, an infestation of the oriental fruit fly on Rota (a 3-sq-mile island) was eliminated by the combination of an attractant and the insecticide naled.

TABLE 29.17 Insect Pheromones

Name	*Structure*	*Species Attracted*
Looplure		Cabbage looper
Muscalure		House fly
Codlelure		Codling moth

With the sex attractants, another approach, called the confusion technique, has been employed. The sex attractant saturates a given area, and the opposite sex is confused by the ubiquity of the signal and fails to find its mate. In the American Southwest, this technique has been employed for the control of the pink bollworm (*Pectinophora gossypiella*). Both qualified successes and failure with this approach have been reported.

The main attractiveness of the pheromone approach is that the biocides need not be broadcast on the crop or soil. In present-day monoculture (i.e., where very large areas are mainly planted in one crop, such as the corn belt in the Midwest) the problem of using pheromones remains to be solved. Insect pheromones are highly specific, but in agriculture the farmer generally is confronted with a complex of insect pests—only rarely a single species. The pheromone approach then must be supplemented with other strategies in order to solve the complex agricultural problem. Although the sex pheromones of many economically important insects have been isolated and their structures elaborated, recent research points to somewhat greater complexity regarding the signaling systems of the insects under study than was originally indicated. Nevertheless, the area of pheromone research provides a broad field for fundamental chemical and field biological exploration. Pheromones have found *practical use* in the determination of the presence and populations of harmful insect species. These monitoring traps are baited with the appropriate pheromone and placed where harmful insect populations might be expected. This monitoring of insect populations eliminates unnecessary field application of toxicants and maximizes their effectiveness when populations genuinely require their use.

Repellents and Antifeeding Compounds

Deet is used as a mosquito repellent, but no repellents have been successfully employed in large-scale agriculture to repel field insects from their normal food supply. Many chemicals exhibit varying degrees of laboatory repellent activity.

Another effective means of insect population management involves the release of predator and/or parasitic insects. There are companies that breed these species and provide them for field usage for crop protection. Certain wasps, lace-wings, ladybugs, predator mites, and other insects suppress but do not eliminate the undesirable species. The practical problems, including the economics of such control, are currently being investigated by government agencies and by certain pest-control companies.

The increase in world trade of agricultural products has resulted in the introduction of nonindigenous and highly destructive insects into environments where such pests have no natural predators. Such infestations (the Oriental fruit fly is an example) are eliminated by the employment of chemical toxicants and by the introduction into a given area of large populations of sterile males. The sterile population is derived from appropriate radiation treatment (gamma rays) or through the use of chemical insect sterilants. In confined areas the release of large populations of sterile males has eliminated the target infestation. After mating, the female produces no viable offspring.

Great enthusiasm existed a few years ago for the development of chemicals that caused sterility in insects. Unfortunately, most of the active chemicals, though quite effective, were powerful alkylating agents, and as such exhibited carcinogenic potential. Some very simple phosphorous amides containing the ethyleneimine $\begin{array}{c} CH_2—CH_2 \\ \diagdown \quad \diagup \\ N \\ | \end{array}$ group were among the most active insect sterilants. $P(O)A_3$ and $P(S)A_3$, where A is the ethyleneimine group, are such agents. Even the solvent hexamethylphosphoric triamide $[[(CH_3)_2N]_3PO]$ possesses some sterilizing capacity. Newer compositions based upon *hormonal* control lead to sterility and are beginning to penetrate the agricultural market.

Rodenticides

Damage to agricultural crops by rodents and the capacity of this group of mammals to harbor disease vectors has stimulated research

on rodenticides. Effective control of various rodent species is achieved through impregnating grain with warfarin. This coumarin compound is related to the chemicals used in medicine to prevent blood clotting, particularly in cardiac disease. As a rodenticide, it causes internal hemorrhage and eventual death.

In addition to the coumarins, certain indanediones, ureas, and other organics find application as rodenticides. Phosphine formed by the slow hydrolysis of metal (usually aluminium and zinc) phosphides continued to be employed for rodent control despite the obvious toxicological hazard. In many third world countries, rodents are a major problem, limiting both the harvest and the storage of grains, as well as being an important public health hazard.

Sulfur and Inorganic Sulfur Compounds

The oldest of the inorganics, sulfur and the water-soluble polysulfides, continue to be employed, and their use may increase during the next decade in many agricultural areas. sulfur, usually as a coarse powder or a small granule, is a soil additive. It is a required nutrient for all plant life, and the soils in many parts of the United States and of the world are deficient in it. The oxidation of sulfur by soil microorganisms assists in the production of acid bodies that help solubilize other insoluble nutrients.

Sulfur is used as a wettable powder or dust, alone or in combination with other pesticides, for foliar fungal control of mildews and other organisms. Calcium polysulfides (CaS_x, $x = 4.5$) usually is employed as a 30 percent aqueous solution in many commercial preparations that are used for soil conditioning and fungal, mite, and insect control. The ammonium and sodium variants also are used.

Because sulfur is a low-cost chemical in abundant supply, an essential nutrient, relatively nontoxic to humans, and ecologically desirable (except in gaseous forms as an air pollutant), its agricultural usefulness may expand. Both industry and agriculture will continue to compete for the sulfur that is removed from fossil fuels (coal, oil, gas).

Inorganic Agrochemicals

Though synthetic organic pesticides dominate the market in countries with advanced agricultural technology, the inorganics still find useful application in several areas. Copper salts, particularly the basic sulfate (Bordeaux mixture) as well as copper oxychloride, control many foliar fungal infections. In Europe, copper products continue to dominate in wine grape mildew-control programs, particularly for the *vin ordinaire* production in the Latin countries.

Basic sulfates of zinc and manganese are used as minor-element nutrient sprays as well as for fungicidal purposes, with or without copper. Basic phosphates also are employed. All of these salts generally are prepared under controlled-pH conditions (sometimes *in situ* in the field) in such a way that a fairly insoluble salt is provided. Thus, its solubility limited the quantity of soluble metal ion available at any given time for absorption by the plant. The high quantities of Zn^{++}, Cu^{++}, Fe^{+++}, and so on, that would be obtained from the normal sulfate, nitrate, or halide sprays would be too phytotoxic for practical use.

In this respect, modern chelation chemistry also is employed to provide metallic minor-element additives. The EDTA (ethylenediaminetetraacetic acid) group combined with metal salts is utilized for both soil and foliar application. Generally, however, less expensive chelants are employed. These include polyphosphates and partially refined products derived from the paper and lumber industry that contain metal coordinating groups.

Calcium arsenate and lead arsenates (orthoarsenates and more complex arsenates) still are employed for certain cotton insects and codling moth control in apples, as well as for other insect-control purposes. These uses are dwindling, as are the uses of arsenites for soil sterilization. At present, the residue and toxicological properties of the arsenicals in relation to human toxicity and ecological factors are under review by government, scientific, and administrative bodies. The above is true also for organic arsenicals such as cacodylic acid $[(CH_3)_2As(O)OH]$ and

sodium methyl arsonate $[NaO)_2As(O)CH_3]$, which are used as herbicides and defoliants.

One of the newer systemic fungicides, foestyl, is the aluminium salt of ethyl phosphite and has activity against diseases caused by Oomycetes such as *Phytophthora* and downy mildews. It is registered for use on nonbearing citrus, ornamentals, and pineapples.

AGROCHEMICAL MANUFACTURE

As the foregoing tables have indicated, there is a wide variety of chemistry involved in the commercial manufacture of agrochemicals. Each material has its own particular route or routes that have been developed by the manufacturing company. Many can be prepared by processes of only a few steps. Other compounds are more complex and require several steps in their manufacture. Obviously the manufacture of the synthetic pyrethroids involves complex chemistry, including stereoselective synthetic procedures or separation and recycling of isomers.

Most manufacturing processes are closely guarded trade secrets. For many agrochemicals several different routes could be envisioned to prepare the desired compound. A variety of considerations affect the decision on which process to use for the manufacture of the desired compound. The route that the chemist used in order to prepare the first sample for testing or the method disclosed in patents may not at all resemble the commercial method. Such factors as raw material availability and cost, plant construction costs, materials of construction, safety factors, by-product recycling, waste disposal, and solvent recovery are among the many variables that affect the choice of how and where to manufacture the desired agrochemical. With time these factors may change and may necessitate a change in the manufacturing process.

Some commonly used starting materials now are controlled substances, among them such materials as methylamine, diethyl malonate, piperidine, ethylamine, pyrrolidine, anthranilic acid, and morpholine. In selecting a synthesis route, it may be advisable to consider using other compounds than these as starting materials. Often dimethyl malonate can be used just as well as diethyl malonate. A derivative of anthranilic acid may work just as well as anthranilic acid.

As many agrochemical intermediates, by-products, and products are corrosive, toxic, flammable, and so forth, they require special materials of construction for a manufacturing plant. Some of the hardest-to-handle intermediates are routinely used in many manufacturing operations, including such materials as chlorine, bromine, hydrogen chloride, sodium, phosgene, phosphorus trichloride, phosphorus oxychloride, carbon disulfide, and many others.

Many agrochemials and their intermediates have toxicological characteristics that necessitate processes and procedures that eliminate exposure of the workers who make or use the toxic materials. Decontamination procedures, fume scrubbing, waste disposal procedures, and reaction conditions need to be carefully considered to make such processes safe for the employees, the environment, and people who live near the plant. Special detection devices, automated processes, special reactors, and so on, may be needed to carry out the reactions safely. Disposal of by-products is becoming an important consideration. The capital costs of new plants are increasing in order to meet these safety and environmental concerns. The day-to-day operating costs of these agrochemical plants also reflect the costs related to human and environmental hazards. Waste disposal facilities may be an important consideration in the location of a plant. Even with the very best of technology, it often is difficult to find a suitable location for a new plant.

Figure 29.2 shows a typical pilot plant for production of field trial compounds, Fig. 29.3 shows a multipurpose small-scale production facility, and a typical flow diagram for the preparation of phorate is presented in Fig. 29.4. Methods of preparation for many different agrochemicals can be found in the references.

Fig. 29.3. Multipurpose small-scale production facility for agrochemicals. (*Courtesy of ICI Americas, Inc.*).

Fig. 29.4. Production and waste schematic for phorate. (*Courtesy of Noyes Data Corporation. Scheme adapted from* Pesticide Manufacturing and Toxic Materials Control Encyclopedia, *p. 612, 1980*).[29]

Captan can be used to illustrate the method of preparation of a typical agrochemical. The reaction steps are as follows:

Reaction 1: The preparation of captan begins with the Diels-Alder addition of butadiene to maleic anhydride to form the 1,2,3,6-tetrahydrophthalic anhydride.

Butadiene Maleic anhydride 1,2,3,6-Tetrahydro-phthalic anhydride

Reaction 2: The anhydride is reacted with ammonia to produce the imide.

1,2,3,6-Tetrahydro-phthalic anhydride Ammonia

1,2,3,6-Tetrahydro-phthalimide Water

Reaction 3: The imide is dissolved in cold caustic and reacted with perchloromethyl mercaptan to give captan. The perchloromethyl mercaptan is prepared by chlorination of carbon disulfide under conditions that eliminate exposure to iron and aluminium.

1,2,3,6-Tetrahydro-phthalimide Perchloromethyl mercaptan

Captan

Another preparation example is that of carbaryl:

Reaction 1: Carbaryl is prepared by reaction of 1-naphthol with methyl isocyanate. The reaction is catalyzed by a tertiary amine, and it is run in a neutral solvent. In order to avoid phytotoxic impurities, the 1-naphthol must be free of the 2-naphthol isomer.

1-Naphthol Methyl isocyanate

Carbaryl

Phorate is an example of the preparation of an organophosphate insecticide:

Reaction 1: The diethyl dithiophosphoric acid is prepared by reaction of phosphorus pentasulfide with ethyl alcohol. (The flow diagram is shown in Fig. 29.4.)

Diethyl dithiophosphoric Formaldehyde Ethyl mercaptan

Phorate

FORMULATION OF AGROCHEMICALS

The art and science employed in transformation of a manufactured chemical into a form that the farmer or the applicator can readily use

in the field is of particular significance to the pesticide industry. This transformation process of chemical to agricultural product is called formulation. At ambient temperatures the technical product is generally either solid (micro- or macrocrystalline or superficially amorphous) or a liquid of variable viscosity. Certain fumigants are rare exceptions to the above. Frequently manufacturing requirements present the formulator with a solution of the agrochemical.

If it is a solid (triazines, many other herbicides, certain chlorocarbons, acid salts, etc.), the problem is to alter the lumps so that $\frac{1}{4}$ to 1 lb of product can effectively cover the foliage on an acre of farm land.

The solid active ingredients may require aging, or hardening, of the microcrystals, followed by comminution and fracturing of the lumps into small particles—generally micro-sized (1—100 micron) particles. This entails a large expenditure of energy and complex engineering equipment. Among the techniques employed are air grinding, and less frequently, wet grinding.

The formulation usually requires inert additives such as clay, celite, talc, and so on, to absorb sticky particles and help prevent adhesion in the final product. If the micronized solid is to be used with water—that is, if it is to be suspended in water and sprayed on the foliage—then surfactants of various types are added. Such a formulation of finely ground particles is called a wettable powder. Frequently a combination of surfactants is employed, including those that wet the solid pesticide, maintain a stable suspension of particles in the water, and help in the adhesion of the suspended pesticide to the foliage surface. Hence the formulation art is replete with references to suspending agents, stickers, spreaders, antiflocculants, and so forth. Because pesticides are used on substances that will become either raw agricultural products or, ultimately, food products, the surfactants, additives, adjuvants, and so on, must all be substances approved by the regulating agencies. In other words, the most efficient surfactant may or may not be acceptable for agricultural formulation. Toxicological and residue factors must be considered for the additives as well as for the pesticides themselves.

Liquid agrochemicals also can be transformed into wettable powders. This is generally done by dispersing the liquid onto the fine particles of an absorbent material such as clay, celite, and others. Again, surfactants and other adjuvants are added to provide the physical and chemical properties necessary for optimum dispersibility in water, wetting of foliage and insect, and so on.

Although used less in recent years, "dust" formulations still are employed in significant amounts. Solids can be ground into fine particles, mixed with inerts, such as talc and clay, and sometimes with adjuvants, and dusted on the plant surface or the soil.

Many farm chemicals are liquids, and many of the solid ones preferably are used as solutions. An emulsifiable concentrate is a formulation of a pesticide that, when mixed with water, provides a dispersion or emulsion containing the active material. Here again, the choice of solvent and surfactant systems involves questions of art and of knowledge of colloid and surface chemitry, as well as the toxicology and residue characteristics of the additives. In the formulation of petroleum oil sprays, heavy deposits with little runoff may be required. Formulation requires a compromise, then, between effective dispersion of the emulsified product in the water during the spraying operation and quick breaking of the emulsion upon the plant surface. In other cases, high surfactant activity is desired to perform the biological objective. Hence, for a single chemical the farmer may require many types of different formulations. Each formulation requires a separate label approved by federal and state authorities.

Nonionic, anionic, and cationic surfactants are employed in agricultural formulations. Although oil in water emulsions are most commonly employed (i.e., the compound is dispersed in a continuous aqueous phase), the reserve (called an invert emulson) is also, but less frequently, used.

Another variation of liquid formulations of agrochemicals is called flowable formulations. Certain plant parts are variably sensitive

to organic solvents, particularly aromatic compounds; and the aliphatic solvents are generally inferior as solvents. Solid farm chemicals can be ground to micron-sized dimensions and suspended at high concentration in a aqueous medium. The adjuvants include surfactants, viscosity-controlling agents, and so on, which enable the formulator to make a useful metastable flowable containing a minimum of water and a high concentration (frequently 50% or over) of solid toxicant. This flowable then is poured by the applicator in the field into a spray tank for dilution and direct application to the crop. There of course is an economic and environmental advantage deriving from the elimination of the solvent.

Dry flowables is the designation given to a rapidly expanding formulation strategy. By wet grinding to dimensions of about 2 microns and the addition of selected adjuvants, formulations with as much as 90 percent active ingredient can be prepared. Wet flowables are at best metastable, whereas dry flowables do not separate into phase layers and require no solvent. Because many of the better solvents are potentially toxic and environmentally undesirable as well, formulation without a solvent is a real advantage. Also paper or cardboard packaging can be employed, thereby reducing the use of nondegradable plastics or metal containers. (*Note*: Useful solvents such as benzene, toluene, xylene, isophorone, dioxane, chloroform, perchloroethylene, ethylene dichloride, methylisobutyl ketone, cyclohexanone, and others are currently suspect or totally interdicted for formulation purposes.) The dry flowables are added directly to the spray tank for surface or air application.

A formulation technique developed during the 1960s uses low-volume and ultralow-volume sprays, particularly of liquid formulations. These formulations, developed to reduce application costs, contain concentrated agrochemical with little or no solvent or additives. Mechanical equipment is used to form liquid droplets of the active material with fairly uniform size (40–100 microns). There is evidence that such formulations increase the biological effectiveness of insecticides (e.g., malation and naled). The problem is one of providing uniform distribution; very fine, submicron droplets will drift past the objective, and large droplets are most inefficient, for many of them never reach the objective. Air generally is used for dispersion. Ground equipment, planes, and helicopters fitted with calibrated delivery systems are commonly used.

A variation of this technique, which is applicable to ground equipment, consists of the vaporization and rapid condensation of an insecticide solution to create a fog, which then is carried by air currents to its objective. In this case, the insect respires submiron particles of the insecticide.

Despite advantages of economic and biological effectiveness there are hazards involved with the use of such low-volume sprays. The fine droplets drift and, because of their biological effectiveness to targets and nontargets, present toxicological and environmental hazards. Therefore, the use of such formulations is on the decline.

Another type of product finding increased use is granule formulations; and one area of their usefulness is the control of disease vectors, particularly mosquito control, in swampy areas. Such targets may be practically inaccessible except by air; but conventional formulations employing droplets or fine solids are of little use, as they will deposit on the foliage surrounding the watery areas in which the mosquitoes breed. Granules of a solid (compacted clay, organic materials, etc.) impregnated with the toxicant can be dispersed by air; they will penetrate through the foliage, and many of them will drop into the surface waters. These granules can be made to release the toxicant at various rates. Granules of various types now are employed in rice culture and in soil-insecticide formulations in the corn belt. Their use is increasing as new applications are found.

Microencapsulation, a formulation technique much used in the pharmaceutical industry, is beginning to have limited application for farm chemicals. It consists of surrounding the toxicant with a polymeric membrane. In one example of this technique, known as interfacial encapsulation, a compatible oil-soluble toxicant

(e.g., methyl parathion, Diazinon, phonophos, etc.) and an oil-soluble monomer are placed in an aqueous environment under agitation and dispersion, and a water-soluble monomer is added to the system. A membrane is formed, encapsulating the toxicant (e.g., methyl parathion) and reducing the toxicological hazard. In the case of methoprene, the polymer reduces degradation of the active ingredient caused photochemically or hydrolytically. The advantages of such formulation must be balanced against the additional cost of the processing; so although encapsulation accomplishes certain objectives, only a small agrochemical market sector has yet been achieved by this technique. (This contrasts with the situation in medicine, where the active agent generally is an expensive ingredient, and the additional formulation cost is a minor consideration.)

Other types of formulations may be employed. Almost all cotton, corn, and wheat seed and many other crop seeds are treated with special preparations that generally contain materials to combat insect and disease attacks, but also may contain growth stimulants, nutrients, and other components. Not only must the additives employed be effective against the intended target; they also must be tolerated by the seed and the emerging seedling under a wide variety of external conditions including variations of soil type, temperature, pH, and moisture. For solid active ingredients in seed treater formulations, a finer particle size than is used in conventional formulations is desired. Such seed must be easily distinguished from untreated seed so that the toxicants employed in the treatment are prevented from entering into the animal and human food chain. Legislation requires the coloring of the treated seed for ready identification, and such treated seed is prohibited from any but its intended use. Nevertheless, tragedies have occurred through human ingestion of seed containing organophosphates, mercurials, and other toxicants employed in the formulation of seed treaters.

As labor in modern American agriculture provides an increasing increment of the total cost, farmers try to reduce the number of applications of the formulations employed. To accomplish this, they sometimes combine fertilizer and pesticide applications but then special formulation problems may arise. Most pesticides are covalent bonded—nonionic organic entities—and the surfactants used with them function best in aqueous solutions of low ionic strength, whereas fertilizers are generally water-soluble ionic compounds. Thus a problem of compatibility must be solved. Novel surfactants that will function in solutions of relatively high ionic strength must be found for combined fertilizer and pesticide formulation to be successful.

Aerosol formulations, dilute pesticide formulations containing a volatile component (usually a fluoromethane, carbon dioxide, propane, or nitrous oxide) that propels the pesticide in fine droplets, are widely used in the home and garden. However, apprehension about ozone depletion in the stratosphere has led to the interdiction of many of the fluorocarbons; so current aerosols employ largely carbon dioxide and propane.

The aim of this presentation has not been to cover the entire area of formulation, but rather to emphasize that formulation frequently is the key to successful application of any agrochemical. The formulation process requires the application of physical-chemical principles, engineering skills, and familiarity with the agricultural problem as a whole. Factors such as corrosion, solubility, surface activity, phytotoxicity, odor, vapor pressure, flammability, compatibility with additives and other formulations, stability, toxicity, and residue characteristics are all part of this art and science.

The formulation of *Bacillus thuringiensis* (BT), mycoherbicides, complex peptides, viruses, and so forth, will require the development of novel techniques, which are discussed in the final sections of this chapter.

BIOTECHNOLOGY AND THE AGROCHEMICAL INDUSTRY

In past editions of this *Handbook*, the chapter on the pesticide industry was almost totally

concenred with synthetic organic, inorganic, and metallo organic compositions with slight references to products derived from plants (extracts, ground plant parts) and a brief mention of some derived from microorganisms.

Today, however, the R&D divisions of major agrochemical companies are devoting increasing proportions of their time, personnel, and dollars to biotechnology. It has been estimated that some major companies (e.g., Monsanto, Du Pont) spend 20 percent or more of their total budgets on biotechnological research. Capital investments also are high. The new biotechnology center at Monsanto is said to have cost $150,000,000. All agrochemical companies participate in varying degrees in R&D in this area.

In the broadest sense, the term "biotechnology" embraces many disciplines, among them fermentation technology, enzymology, protein preparation and separation, microbiology, cell culture, the transformation of DNA and genetic materials, and others.

In this chapter mention has been made of marketable products such as the avermectins for mite, insect, and animal parasite control and *Bacillus thuringiensis* for control of various insect species, which are triumphs of the application of fermentation and microbiological applied science. Many new products are currently under development in this area. In addition, amino acids and other chemical intermediates, particularly those with chiral centers, now are produced economically with such techniques.

The beginning of the twentieth century witnessed the improvement of plant species by classical genetic improvement. Modern molecular biology, which encompasses the alteration of DNA through the use of restriction enzymes, has become the most exciting and most publicized area for academic and industrial research and development.

The agrochemical industries' R&D is heavily involved in improvement of the production rate, quantity, and quality of microorganism-derived products. Examples of this are antibiotics in general and microorganism-derived insecticides (e.g., B.T. with broader insect-killing scope or greater biological activity), mycoherbicides, and fungicides with enhanced activity. For prokaryotes (microorganisms without a nuclear cell wall) considerable development progress is being made.

Much effort is currently devoted to altering the genetics of food and fiber crops. One area is the inclusion of genes in plants that will make them resistant to insect, microorganism, and virus attack. The inclusion of such genes in the plant genome (with repoductive ability) is an order of magnitude more difficult with these eukaryotes (organisms with nuclear cell walls) than with the prokaryotes.

There are other obstacles that such genetic transformations must overcome. There is a public perception in the United States and much of Europe that is hostile to the very concept of this new, human-performed genetic manipulation. Further regulatory mechanisms will have to be developed for plants with, for example, insect-controlling abilities, not heretofore a part of the plants' biochemical constitution. Plants that do possess pest-controlling capacity could and probably will participate in the generation of resistant varieties of predators (insects, fungi, bacterial viruses). These are hurdles that the products derived from gene manipulation must overcome.

The decade of the 1990s will be involved with increased biotechnological R&D. Marketable products will come from genetic alteration—first from prokaryote alterations. Deveopment will be slower for the genetic manipulation of eukaryotes. It is our conviction that in the nineties there will be successes both for the classical synthetic agrochemicals and for certain biotechnological developments.

AGROCHEMICALS FOR THE FUTURE

Biological, Genetic, and Synthetic Strategies

As a result of escalating costs for the development of new synthetic organic insecticides, as well as the increase of resistance and cross resistance to these synthetics and environmental concerns, real and perceptual,

much R&D interest is exhibited in biological alternatives to synthetic organics.

Bacillus thuringiensis (B.T.) is one of the alternative approaches that has advanced to the point of successful marketable products. B.T. is a gram-positive soil inhabitant, various strains of which produce proteins that are broken down in the insect gut into highly insecticidal fragments, 27 to 140 kilo Daltons (KD). The latter may be further protolytically fragmented to insect-toxic polypeptides. Mammalian digestive systems appear to be unaffected by B.T., and its wide toxicity index, insect versus mamma, contributes to the usefulness of this material for crop applications. B.T. itself is highly active against various lepidoptera (e.g., armyworms, cabbage looper, gypsy moth, and spruce budworm) with remarkable specificity between strains. Closely related to B.T. and *Bacillus sphaericus* (diptera, mosquitoes), *Bacillus popilliae* (Japanese beetle), and others.

Much current R&D centers on genetic modification of B.T. itself and on the introduction of genes responsible for the insect-toxic proteins into organisms associated with plant growth and into plants themselves. In these strategies the plant-associated genetically altered organisms would protect the plant, or the plant itself would express these protective insect toxins. There are many hurdles to overcome for these new strategies, and none of them has advanced beyond the lab or regulated small plot stage. Nevertheless, both in academia and in the agrochemical industry a substantial portion of research budgets focuses on such efforts. Should the above-mentioned hurdles (resistance implications, regulatory concerns, high yield and stable fermentation products, etc.) be overcome, appreciable markets could develop from this strategy. Estimates for current B.T. sales are about 0.6 percent of the total insecticide market. Seed companies, many of which are either owned by the larger agrochemical companies or have business arrangements with them, are highly interested in these plant modifications.

Genetic engineering also provides opportunities for modification of microorganisms and plants other than for the introduction of insect toxins. R&D targets for plants are fungal, bacteral, and viral protection. Modification of plants to cause them to produce desirable amino acids and peptides, saccharides, oils, and so on, is receiving considerable attention. For these eukaryotes the molecular biology involved is an order of magnitude more difficult than such modification of the prokaryotes. The modification of microorganisms to produce the above types of product is further advanced than such modification of plants or animals. Products from microorganisms, from simple organic intermediates to complex substances such as human insulin, are already on the market.

Another area of current R&D concentration is insect control through insect viral infection. The virus that infects *Heliothis*, the nuclear polyhedrosis virus, can be propagated and has been used with variable success. Other insect virus agents also are being investigated, as is the molecular biology of the virus.

It has been well known that various fungi and bacteria infect important weed species. The Department of Agriculture laboratories and certain companies are pursuing this approach to combat noxious weeds, and are attempting to pioneer products known as mycoherbicides. Another exploratory project involves the use of certain types of nematodes that can infect and ultimately kill insects. These nematodes can be cultured, harvested, and field-distributed for the control of a variety of pests in addition to insects.

In the more classical areas of synthetic organic agrochemicals, there recently has been great success in reducing the environmental burden carried by their agricultural application through reduction of the dose per acre (higher biological activity). Such efforts are continuing. Further, a more biologically rational approach to synthesis is developing. Much progress has been made in understanding basic biochemical processes in plants and microorganisms. Knowledge of the chemical constitution of enzymes, of receptors, and of metabolism in general is opening the door to the bio-rational approach to synthesis. Future agrochemicals will result from the successful

application of this new knowledge to the creation of new and useful compositions.

Developmental Agrochemicals

Agrochemicals are a critical resource in worldwide food production and public health. The growing world population and the expanding development of eastern block and third world countries increasingly require new and more effective agrochemicals. The development of these new materials is a slow and costly process. Tens of millions of dollars and upwards of eight to ten years are required before any new product arrives at the marketplace. Anywhere in this long process, a new material can be dropped from development for a variety of reasons. By the time that a new material comes to public attention, several years already have been spent in the development of the agrochemical. Table 29.18 lists a variety of new materials that currently show promise. Other developmental and experimental compounds were cited earlier.

Future Trends for the Agrochemical Industry

The agrochemical industry will continue to be a dynamic industry into the twenty-first century. The 1980s saw a consolidation of companies. The high cost of registration and development of new agrochemicals in part fueled the drive toward acquisitions and takeovers. Only the very large companies seem able to operate successfully in this climate. The trend toward globalization will continue. Fewer but larger corporate entities will comprise the industry as a whole. These larger companies will have to be flexible in meeting customer needs if they are to survive.

Struggles over real and perceptual hazards due to chemicals in general and agrochemicals in particular will intensify in the 1990s and

TABLE 29.18 Examples of Developmental Agrochemicals

Company	*Code Number*	*Structure*	*Uses*
Du Pont	DPX-E9636		Postemergence herbicide in corn
Ishihara	SL-950		Postemergence herbicide in corn
Schering	SN-106279		Herbicide in winter cereals
Sumitomo	S-23121		Pre- and postemergence herbicide
Rohm and Haas	RH-0265		Postemergence herbicide for broadleaf weeds

TABLE 29.18 Examples of Developmental Agrochemicals (continued)

Company	*Code Number*	*Structure*	*Uses*
Uniroyal	UBI-C4874		Postemergence grass herbicide
Asahi	AKH-7088		Postemergence herbicide for broadleaf weeds in soybeans
Hokko and Rohm and Haas	HOK-1566, RH-0898		Herbicide for use in wheat and rice
CIBA-GEIGY	CGA-163935		Plant growth regulator for small grain cereals, rape, and turf
Rohm and Haas	RH7592		Fungicide for wheat
CIBA-GEIGY	CGA 142705		Seed treatment fungicide
CIBA-GEIGY	CGA 80000		Systemic fungicide
Nissan	NC-170		Juvenile hormone for leaf and plant hoppers
Lilly	LY211795		Powdery mildewicide
Rohm and Haas	RH7988		Systemic aphicide

will be a major subject of political debate and legislation for the United States and much of the world. Agriculture itself, upon which the agrochemical industry is dependent, has been a partially or wholly subsidized industry all over the earth. Nowhere, presently, does a free market exist for the agricultural industry; trade barriers to free imports and exports of agricultural products are the rule. We see no immediate resolution of this problem but believe that in the long run these barriers will be reduced, perhaps removed. Until then the agrochemical industry will have to operate in an uncertain and volatile economic environment.

The use of classical agrochemicals will be one of the strategies for effective and economic pest control. At a reduced rate, newer compositions will be developed and registered; and these new chemicals will on the whole be more highly active than current materials. Their use will reduce human and environmental hazards, but their development and commercialization will require huge investments for the research required to guarantee their safety.

There continues to be an explosion of knowledge on all fronts. Advances in chemistry influence and in some cases generate discoveries in interdisciplinary areas, as they do in agronomy, toxicology, physiology, genetics, and biochemistry. The industry will more fully support academic research in areas where it is needed for an understanding of the effects of chemicals on health and environmental issues.

The complexities of all modern science make computers a necessary tool. Huge data bases are being developed in all the scientific disciplines. Better understanding on a molecular basis of biological processes and metabolism leads to the conviction that some of the newer agrochemicals will arise from the bio-rational extension of this new knowledge, supplementing the classical synthesis and bio-evaluation approach.

The 1990s will be the decade for the trials, scientific and economic evaluations, and development of biotechnological products. The decade will be marked by the successful emergence of new pest-control strategies and biotechnological products at a slow to moderate pace into the beginning of the twenty-first century. Many of the biotechnological venture capital companies will be absorbed by the large corporations of the agrochemical industry. This trend has already begun.

The challenge then to all workers in this field is to understand the complex interrelatonships involved in providing safe and effective control agents to meet world's needs in the twenty-first century.

REFERENCES

Chemical, Biological, and Mode of Action

1. Audus, L. J. (Ed.), *The Physiology and Biochemistry of Herbicides*, Academic Press, New York, 1964.
2. Ashton, F. M., *Mode of Action of Herbicides*, John Wiley & Sons, New York, 1973.
3. Baja, Y. P. S. (Ed.), *Biotechnology in Agriculture and Forestry* (crops, Vol. 1 and Vol. 2) Springer-Verlag, Berlin, Heidelberg, New York, 1986 (Vol. 1), 1988 (Vol. 2).
4. Baker, D. R., Fenyes, J. R., Moberg, W. K., and Cross, B. (Eds.), *Synthesis and Chemistry of Agrochemicals*, ACS Symposium Series, No. 355, American Chemical Society, Washington DC, 1987.
5. Boeger, P., and Sandmann, G. (Eds.), *Target Sites of Herbicide Action*, CRC Press, Boca Raton, FL, 1989.
6. Buechel, K. H. (Ed.), *Chemistry of Pesticides*, John Wiley & Sons, New York, 1983.
7. Coats, J. R. (Ed.), *Insecticide Mode of Action*, Academic Press, New York, London, 1982.
8. Conner, J., Jr., Ebner, L., O'Connor, C., Voz, C., and Weinstein, K. (Eds.), *Pesticide Regulation Handbook*, Executive Enterprise Publications Co., Washington DC, 1987.
9. Corbett, J. R., Wright, K., and Baillie, A. C., *The Biochemical Mode of Action of Pesticides*, 2nd ed., Academic Press, London, 1984.
10. Crafts, A. S., *The Chemistry and Mode of Action of Herbicides*, John Wiley & Sons, New York, 1961.
11. *Crop Protection Chemicals References*, 5th ed., Chemical and Pharmaceutical Press, New York, 1989.

12. Elliott, M. (Ed.), *Synthetic Pyrethroids*, ACS Symposium Series, No. 42, American Chemical Society, Washington DC, 1977.
13. Fest, C., and Schmidt, K. W., *The Chemistry of Organophosphorus Pesticides*, Springer-Verlag, New York, Heidelberg, Berlin, 1973.
14. Fletcher, W. W., and Kirkwood, R. C., *Herbicides and Plant Growth Regulators*, Granada, London, 1982.
15. Georghiou, G. P., and Saito, T. (Eds.), *Pest Resistance to Pesticides*, Plenum Press, New York, London, 1983.
16. Green, M. B., and Spilker, D. A. (Eds.), *Fungicide Chemistry: Advances and Practical Applications*, ACS Symposium Series, No. 304, American Chemical Society, Washington DC, 1986.
17. Hartley, D. H., et al. (Eds.), *The Agrochemicals Handbook*, 2nd ed., Royal Society of Chemistry, Nottingham, 1987.
18. Hedin, P. A. (Ed.), *Bioregulators for Pest Control*, ACS Symposium Series, No. 276, American Chemical Society, Washington DC, 1985.
19. Heitz, J. R., and Downum, K. R. (Eds.), *Light-Activated Pesticides*, ACS Symposium Series, No. 339, American Chemical Society, Washington DC, 1987.
20. Hilton, J. L. (Ed.), *Agricultural Chemicals of the Future*, (Beltsville Symposia in Agricultural Research), Rowman and Allenheld, Totowa, NJ, 1985.
21. Kearny, P. C., and Kaugmann, D. D., *Herbicides: Chemistry, Degradation and Mode of Action*, 2nd ed., 2 Vols., Marcel Decker, New York, 1975.
22. Kohn, G. K. (Ed.), *Mechanism of Pesticide Action*, ACS Symposium Series, No. 2, American Chemical Society, Washington DC, 1974.
23. Mandava, N. B. (Ed.), *Plant Growth Substances*, ACS Symposium Series, No. 111, American Chemical Society, Washington DC, 1979.
24. Magee, P. S., Kohn, G. K., and Menn, J. J. (Eds.), *Pesticide Synthesis through Rational Approaches*, ACS Symposium Series, No. 255, American Chemical Society, Washington DC, 1984.
25. Marco, G. J., Hollingworth, R. M., and Durham, W. (Eds.), *Silent Spring Revisited*, American Chemical Society, Washington DC, 1986.
26. Menn, J. J., and Beroza, M. (Eds.), *Insect Juvenile Hormones: Chemistry and Action*, Academic Press, New York, 1972.
27. Ragsdale, N. N., and Kuhr, R. J. (Eds.), *Pesticides: Minimizing the Risks*, ACS Symposium Series, No. 336, American Chemical Society, Washington DC, 1987.
28. Sine, C., et al. (Eds.), *1990 Farm Chemicals Handbook*, Meister Publishing Co., Willoughby, OH, 1990.
29. Sittig, M. (Ed.), *Pesticide Manufacturing and Toxic Materials Control Encyclopedia*, Noyes Data Corp., Park Ridge, NJ, 1980.
30. Torrey, S. (Ed.), *Preemergence Herbicides*, Noyes Data Corp., Park Ridge, NJ, 1982.
31. U.S. Environmental Protection Agency, *Pesticide Fact Handbook*, Noyes Data Corp., Park Ridge, NJ, 1988.
32. Ware, G. W., *The Pesticide Book*, 3rd ed., Thomson Publications, Fresno, CA, 1989.
33. Worthing, C. R., et al. (Ed.), *The Pesticide Manuel*, 8th ed., The British Crop Protection Council, Thorton Heath, U.K., 1987.

Statistical

34. *Chemical Economics Handbook*, Stanford Research Institute, Palo Alto, CA. (annually).
35. U.S. Tariff Commission, *Synthetic Organic Chemicals, U.S. Production and Sales*, U.S. Government Printing Office, Washington DC. (annual reports).

Agrochemical Journals

36. *Jourhal of Agricultural and Food Chemistry*, Vol. 1 (1953)–current, American Chemical Society, Washington DC.
37. *Pesticide Biochemistry and Physiology*, Vol. 1 (1971)–current, Academic Press, Duluth, MN.
38. *Pesticide Science*, Vol. 1 (1970)–current, Society of Chemical Industry, Elsevier Applied Science, Barking, Essex, U.K.
39. *Residue Reviews*, Vol. 1 (1962)–Vol. 96 (1986), Springer-Verlag, New York.
40. *Reviews of Environmental Contamination and Toxicology*, Vol. 98 (1987)–current, Springer-Verlag, New York.
41. *Weed Research*, Vol. 1 (1960)–current, European Weed Research Society, Blackwell Scientific Publications, Oxford, U.K.
42. *Weed Science*, Vol. 1 (1952)–current, Weed Science Society of America, Champaign, IL.

30

Chemical Explosives

Walter B. Sudweeks*

and

Rocket Propellants

George M. Clark** and Felix M. Chen**

I. CHEMICAL EXPLOSIVES

INTRODUCTION

Without explosives we probably would still be riding around in horse-drawn carriages; yet when most people think of explosives, they think only of bombs and artillery shells used in warfare. It appears, however, that more explosives have been used in industry for peaceful purposes than in all the wars put together.[1,2] According to U.S. Bureau of Mines reports, industrial consumption of explosives in the United States alone totaled nearly 110 billion lb from 1912 through 1989.[3] Industrial explosives are currently being produced in the United States at a rate of nearly 5 billion lb/year.

Historically, the discovery of explosives must be considered one of the greatest milestones in the development of modern society. Although the use of explosives for military and destructive purposes may be viewed as an overall negative for humankind, this application has undeniably had a significant impact on the shaping of territorial boundaries and the stability of the governments of most of the industrialized world. Indeed it has been suggested that the advent of explosives and their military power was instrumental in the fall of the feudal system that permitted the establishment of modern civilization's antecedent forms of government.[4,5] Up to that time power was maintained simply by the strength of physical barriers. However, it is the commercial use of explosives in coal mining, metal mining, and construction that has made possible the dramatic technological advancements of the past 200 years. Through the efficient use of explosive energy people were freed from back-breaking labor, and miners were able to obtain sufficient quantities of iron, copper, silver, nickel, zinc, lead, and so on, from the earth to allow the industrial revolution to take place.

Without explosives the steel industry and our whole transportation system would not be possible. The same is true, of course, for our vast array of household conveniences that contain various metal parts. The generation of electricity has been almost entirely dependent

*IRECO, Inc.: Section on Chemical Explosives.
**Aerojet Propulsion Division: Section on Rocket Propellants.

TABLE 30.1 Examples of Explosive Applications

Coal mining	Explosive welding of pipelines
Metal mining	Explosively created art
Construction	Petroleum fire control (burning tankers and gas wells)
Quarrying	Medicine (heart pills, vasoinflators, kidney stone removal)
Trenching	Primers for bullets and shotgun shells
Demolition of old buildings and bridges	Military bombs and artillery shells
Automobile air bags	Nuclear blast simulation
Geothermal and oil well stimulation	Industrial diamond manufacture
Seismic exploration	Rocket stage separation
Firebreaks	
Explosive bolts	
Metal cladding	
Metal cutting	

upon coal, and coal mining today is still the largest consumer of industrial explosives. Rock quarrying for road and building materials and excavation for skyscrapers, tunnels, roads, pipelines, and utilities are direct beneficiaries of the labor-saving use of explosives.

Table 30.1 lists these and other more esoteric uses of chemical explosives. This compilation is by no means comprehensive, but it may help us to appreciate the wide breadth of explosive applications that exist today.

DEFINITION

A chemical explosive may be defined as a compound or mixture of compounds that reacts very rapidly to produce relatively large amounts of gases and heat. It is the nearly instantaneous formation of gases plus their rapid expansion due to pressure and heat that results in the destructive force or useful work illustrated in Fig. 30.1. Besides chemical explosives, there are nuclear and physical explosives, the latter being of only minor importance. With nuclear explosions it is primarily the instantaneous release of massive amounts of heat that does work by vaporizing and expanding the surrounding medium. Examples of physical explosions are steam bursts, due to the rapid heating and vaporization of water, and the sudden release of

Fig. 30.1. Explosives at work. A typical blast at a New England rock quarry.

pressurized gases. Throughout this chapter the general term "explosives" is used to refer to chemical explosives.

CHEMISTRY OF COMBUSTION

For a simple understanding of explosives it is helpful to compare an explosive reaction with the more familiar combustion or burning reaction. Fire safety is often introduced by reference to a fire triangle, as shown in Fig. 30.2. This geometric figure shows the three elements needed to have a fire: fuel, oxygen, and a source of ignition. The process of combustion is basically an oxidation–reduction (redox) reaction between the fuel and oxygen from the air. Once initiated, this reaction becomes self-sustaining and produces large volumes of gases and heat. The heat given off further expands the gases and provides the stimulus for the reaction to continue.

The burning reaction is a relatively slow process, depending upon how finely divided the fuel is, that is, the intimacy of contact between the fuel and the oxygen in the air. Because burning is diffusion-controlled, the more intimately the fuel and oxygen are mixed, the faster they can react. Obviously, the smaller the particles of fuel, the faster combustion can occur. A wood log burns quite slowly, but the same log slit into kindling will burn quite rapidly. Sawdust and dry wood powder will burn even faster if they are dispersed in air. In this configuration, the burning reaction even can be fast enough to produce an explosion. The sensitivity of a fuel–oxygen mixture, that is, the ease of initiation, also is dependent on the intimacy of the mixture. Therefore, in our example, kindling is much easier to ignite than a log, and a cloud of flour or gasoline fumes in air requires only a spark. In each case the fundamental difference is simply a function of how finely divided the fuel component is and how readily accessible it is to the oxygen.

Fig. 30.2. The fire safety triangle.

Another result of the fineness or particle size of the fuel is the completeness or efficiency of the reaction. In a complete combustion all the fuel elements are oxidized to their highest oxidation state. Thus, wood, being mainly cellulose, and gasoline, being generally a hydrocarbon (e.g., octane), produce primarily CO_2 and H_2O vapor upon complete combustion as shown:

$$\underset{\text{Cellulose}}{-C_6H_{10}O_5-} + 6O_2 \rightarrow 6CO_2 + 5H_2O + \text{heat}$$

$$\underset{\text{Octane}}{C_8H_{18}} + 12.5O_2 \rightarrow 8CO_2 + 9H_2O + \text{heat}$$

Once initiated these burning reactions give off heat energy, which sustains the reactions. Heat is released because the oxidized products of the reaction are lower in energy (more stable) than the reactants. The maximum potential energy release can be calculated from the respective heats of formation of the products and reactants. Actual heats of combustion can be measured experimentally by causing the reaction to occur in a bomb calorimeter. The calculated energy values for the above reactions are -3857 cal/g of cellulose and $-10{,}704$ cal/g of octane, respectively.

In the case of an inefficient burn, some less stable or higher-energy products are formed so that the resultant heat energy given off is lower than that for complete combustion. In the above examples inefficient combustion could result for lack of oxygen accessibility, producing carbon monoxide or even carbon particles instead of carbon dioxide. A smoky flame is evidence of unburned carbon particles

Fig. 30.3. An explosion triangle.

and results from inefficient combustion where fuel particles are so large or so dense that oxygen cannot diffuse to the burning surface fast enough. If this inefficiency is great enough, insufficient heat is given off to keep the reaction going, and the fire will die out.

All explosive reactions involve similar redox reactions; so the above principles of combustion can help illustrate, in a very basic way, the chemistry involved in explosions. As in a fire, three elements (fuel, oxidizer, ignition source) are needed for an explosion. Figure 30.3 shows an explosion triangle, which is similar to the fire triangle. In general, the products of an explosion are gases and heat although some solid oxidation products may be produced, depending upon the explosive composition. As in normal combustion, the gases produced usually include carbon dioxide and water vapor plus other gases such as nitrogen, again depending upon the composition of the explosive.

Figure 30.3 differs from Fig. 30.2 in two very significant ways. First, oxygen from the air is not a major reactant in the redox reactions of most explosives. The source of oxygen (or other reducible species) needed for reaction with the fuel—the oxidizer—may be part of the same molecule as the fuel or a separate intermixed material. Thus an explosive may be thought of as merely an intimate mixture of oxidizer and fuel. This degree of intimacy contributes to the second significant difference between an explosion and normal combustion —the speed with which the reaction occurs.

Explosives in which the oxidizer and fuel portions are part of the same molecule are called molecular explosives. Classical examples of molecular explosives are TNT (2,4,6-trinitrotoluene), PETN (pentaerythritol tetranitrate), and NG (nitroglycerin or, more precisely glycerol trinitrate). The chemical structures of these explosives are shown in Fig. 30.4. As can be seen in the structures, the oxidizer portions of the explosives are the nitro ($—NO_2$) groups in TNT and the nitrate ($—ONO_2$) groups in PETN and NG. The fuel portions of all three explosives are the carbon and hydrogen (C and H) atoms. Comparison of the ratios of carbon to oxygen in these explosives (i.e., approximately 1:1 for TNT, approximately 1:2 for PETN, and 1:3 for NG) shows that TNT and PETN are deficient in oxygen; that is, there is insufficient oxygen present in the molecule to fully oxidize the carbon and hydrogen. Consequently, products such as carbon monoxide, solid carbon (soot), and hydrogen are produced, as well as carbon dioxide and water vapor. Prediction of the exact products of explosion is complex, especially for the oxygen-deficient

Fig. 30.4. Chemical structure of three molecular explosives.

explosives, because the ratios of CO_2, CO, H_2O, and H_2 will vary, depending upon reaction conditions (explosive density, degree of confinement of the explosive, etc.).[6,7] The following equations show typical ideal reaction products along with calculated heats of reaction for these molecular explosives:

$$\underset{\text{TNT}}{C_7H_5N_3O_6} \rightarrow 1.5CO_2 + 0.5CO + 2.5H_2O + 1.5N_2 + 5C + 1290\,\text{cal/g}$$

$$\underset{\text{PETN}}{C_5H_8N_4O_{12}} \rightarrow 4CO_2 + 4H_2O + 2N_2 + C + 1510\,\text{cal/g}$$

$$\underset{\text{NG}}{C_3H_5N_3O_9} \rightarrow 3CO_2 + 2.5H_2O + 1.5N_2 + 0.25O_2 + 1480\,\text{cal/g}$$

Explosives in which the oxidizer and fuel portions come from different molecules are called composite explosives because they are a mixture of two or more chemicals. A classic industrial example is a mixture of solid ammonium nitrate (AN) and liquid fuel oil (FO). The common designation for this explosive is the acronym, ANFO. The oil used (typically #2 diesel fuel) is added in sufficient quantity to react with the available oxygen from the nitrate portion of AN. The redox reaction of ANFO is as follows:

$$\underset{\text{AN}}{3NH_4NO_3} + \underset{\text{FO}}{-CH_2-} \rightarrow CO_2 + 7H_2O + 3N_2 + 880\,\text{cal/g}$$

"Oxygen balance" is the term applied to quantify either the excess oxygen in an explosive compound or mixture (beyond what is needed for complete combustion of the fuel elements) or oxygen deficiency (compared to the amount required for complete combustion). It is expressed as either a percentage or a decimal fraction of the molecular weight of the oxygen in excess (+) or deficiency (−) divided by the molecular weight of the explosive or the ingredient being considered. Individual components of an explosive mixture have oxygen balance (O.B.) values that may be summed for the mixture. Shown below are the O.B. calculations for ammonium nitrate and fuel oil:

$$\underset{\substack{\text{AN}\\ \text{Mol. wt.} = 80}}{NH_4NO_3} \rightarrow 2H_2O + N_2 + \underset{\text{Mol. wt.} = 32}{\tfrac{1}{2}O_2}$$

$$\text{O.B.} = \frac{+(1/2)(32)}{80} = +0.20$$

$$\underset{\substack{\text{FO}\\ \text{Mol. wt.} = \sim 14n}}{\text{-}(CH_2)_n\text{-}} + \frac{3n}{2}O_2 \rightarrow nCO_2 + nH_2O$$

$$\text{O.B.} = \frac{-(3n/2)(32)}{14n} = -3.43$$

From the O.B. values, one can readily determine the ratio of ingredients to give a zero oxygen balance mixture for optimum efficiency and energy. Thus the weight ratio for ANFO is 94.5 parts of AN and 5.5 parts of FO ($94.5 \times 0.20 = 5.5 \times 3.43$).

For the molecular explosives shown previously, the respective oxygen balances are: TNT, −0.74; PETN, −0.10; and NG, +0.04. Thus, nitroglycerin is nearly perfectly oxygen-balanced; PETN is only slightly negative; but TNT is very negative, meaning significantly deficient in oxygen. Therefore, combinations of TNT and ammonium nitrate have been employed to provide additional oxygen for the excess fuel, as, for example, in the Amatols developed by the British in World War I.[8]

The speed of combustion in a normal fire is relatively slow, on the order of a few inches per hour to a few inches per second, whereas the realm of explosion typically covers reaction speeds of around 3000 ft/sec to 5 or 6 miles/sec. As discussed with the fire triangle above, the speed of compustion is determined by the intimacy of contact between the fuel and oxygen. With sufficient intimacy of mixing, combustion can reach explosive speeds. In fact, early in this century considerable industrial usage of liquid oxygen explosives (LOX) occurred.[9] These were generally bags of charcoal that were soaked for a couple of hours in liquid oxygen and then

quickly loaded into boreholes and initiated before too much oxygen evaporated.[10]

Modern commercial explosives react in a very rapid and characteristic manner referred to as a detonation. Detonation has been defined as a process in which a shock-induced supersonic combustion wave propagates through a reactive mixture or compound.[11] This high pressure shock wave compresses the reactive material in contact with it so that rapid heating of the material, initiation of chemical reaction, and liberation of energy result. This energy, in turn, continues to drive the shock wave. Pressure in a detonation shock wave may reach millions of pounds per square inch. The sudden pressure pulse shocks the explosive material as it passes through, causing a nearly instantaneous chemical reaction in the body of the explosive. Once initiated, molecular explosives tend to reach a steady-state reaction with a characteristic detonation velocity. Composite or mixture explosives also have steady-state detonation velocities, but these velocities are more variable than those of molecular explosives and are influenced by such factors as diameter of the charge, temperature, and confinement.

HISTORICAL DEVELOPMENT

The first known explosive material was black powder, a mixture of potassium nitrate (saltpeter), charcoal, and sulfur. As such it is a composite explosive whose properties are dependent upon how finely divided each of the ingredients is, and how intimately they are mixed. The exact origins of black powder are lost in antiquity. Publications referring to it seem about equally divided between those that attribute its origin to third- or fourth-century China[12–15] and those that place it closer to the thirteenth century, at about the time of Roger Bacon's written description in 1242.[4,16–19] Nevertheless, its use did not become very popular until the invention of the gun by Berthol Schwartz in the early 1300s; and its first recorded use in mining did not occur for over 300 years after that. Following its introduction for blasting in 1627, the use of black powder for mining, road building, and quarrying spread rapidly as a tremendous labor-saving invention. Before then rock breaking could be done only by pick and shovel or mechanical means such as fire setting, that is, heating the rock with a fire and then pouring water on it to cause cracking.

The formulation for black powder has varied somewhat in the ratio of ingredients over the years, but it has generally stabilized at its present 75 percent nitrate, 15 percent charcoal, and 10 percent sulfur.[12] In 1857 Du Pont replaced potassium nitrate with the more readily available sodium nitrate and found comparable performance.[20] The production and application of black powder played a critical role in the rapid expansion of the United States in the early nineteenth century as canals were dug and railroads built to span the continent.

For over 200 years black powder was the only blasting agent known, but the 1800s brought a number of rapid developments that would lead to its demise and replace it with safer and more powerful explosives. Table 30.2 presents a chronological summary of some of the significant discoveries of the 1800s. Credit for the first preparation of nitroglycerin is generally ascribed to Ascanio Sobrero in Italy in 1846. In that same year C. F. Schönbein, a German physicist, recognized the explosive potential of nitrocellulose, which had been prepared a few years before. Sobrero advised against the use of NG because of its hazardous

TABLE 30.2 Nineteenth-Century Explosive Discoveries

1800	Mercury fulminate
1846	Nitrocellulose
1846	Nitroglycerin
1847	Hexanitromannite
1862	Commercial production of nitroglycerin
1867	Dynamite
1867	Blasting cap
1867	Ammonium nitrate explosive patented
1875	Blasting gelatin and gelatin dynamite
1884	Smokeless powder
1886	Picric acid
1891	TNT
1894	PETN

and unpredictable nature, but the Swedish inventors Emmanuel Nobel and his son Alfred took an interest in this powerful liquid explosive and produced it commercially in 1862. At first the liquid was poured directly into the holes bored in rocks and was ignited by a trail of black powder. However, its transportation and its handling were very hazardous, and eventually Alfred Nobel discovered that NG that had accidentally soaked into a granular type of packing material was still explosive, but was much safer to handle than the liquid. He called this new invention "dynamite." In its partially desensitized state, however, dynamite was difficult to ignite by the usual methods used for pure NG. Therefore, also in 1867, Alfred Nobel devised the blasting cap using mercury fulminate, a very sensitive explosive that had been discovered at the beginning of the century. With this development dynamite became the foundation of the commercial explosives industry, and its name was soon adopted into every major language of the civilized world.

With its higher detonation velocity, dynamite had a much higher brisance or shattering effect than black powder and was more effective at breaking hard rock and ore formations. Its rapid rise in popularity resulted in fierce copycat competition internationally. While searching for an improved, new product to keep ahead of competitors, Nobel found that a combination of NG and nitrocellulose produced a gel with even more explosive power and better water resistance than contemporary dynamite formulations. In 1875 he received patents for both blasting gelatin and gelatin dynamite. The success of these products was largely responsible for Nobel's later wealth, with which he eventually established the Nobel prizes.

For military and gun applications black powder continued to be the only explosive of choice as a propellant or bursting charge until the inventions of the late 1800s, when smokeless powder, based on nitrocellulose, proved to be a cleaner, safer, and more effective propellant than black powder. The synthesis of picric acid (2,4,6-trinitrophenol) followed by TNT and PETN gave solid, powerful, molecular explosives of more uniform performance for use in bombs and artillery shells. Picric acid and TNT are relatively low-melting solids that could be melt-cast as bomb fills. Mixtures of TNT and PETN (known as Pentolite) became useful as boosters for larger, less sensitive explosive charges in both military and industrial applications. The main explosives used in World War I were TNT, Tetryl (2,4,6-trinitrophenylmethylnitramine), and Hexyl (hexanitrodiphenylamine), and in World War II they were TNT, PETN, and RDX (1,3,5-trinitro-1,3,5-triazacyclohexane).[21]

In the industrial arena the production of black powder in the United States dropped precipitously after reaching a peak of 277 million lb in 1917.[20] By the mid-1960s it had ceased to be of commercial significance, but during the same time period dynamite production rose from 300 million lb to 600 million lb.

In 1947 a spectacular accident of catastrophic proportions ushered in the next revolution in explosives. Following the end of World War II, excess nitric acid capacity (previously used to manufacture TNT and other explosives for the war effort) was converted to making ammonium nitrate fertilizer for the agricultural industry. Solid ammonium nitrate produced in the form of spherical particles called prills was coated with paraffin wax to prevent caking, loaded into 100-lb paper bags, and shipped to Texas City, Texas, where it was loaded into ships in Galveston Bay. On the morning of April 16, soon after loading was resumed aboard the SS *Grandcamp*, smoke was observed coming from one of the holds. Along with other cargo the partially loaded ship contained 2300 tons of fertilizer-grade ammonium nitrate. Upon investigation, a fire was discovered in one of the holds containing ammonium nitrate. Limited efforts to extinguish the fire were unsuccessful, and an hour later the bulk of the coated fertilizer detonated, killing 600 people and injuring 3000.[22]

This tragedy, along with several other large-scale accidents involving ammonium nitrate explosions, finally led researchers

to the conclusion that inexpensive, readily available, fertilizer-grade ammonium nitrate could be used as the basis for modern industrial explosives. Actually ammonium nitrate had been used for years as an additive to dynamites to increase their strength or lower the amount of nitroglycerin required, and the canned Nitramon blasting agents utilized in the 1930s and 1940s were mainly ammonium nitrate. In fact, explosive formulations consisting of ammonium nitrate and various fuels were patented contemporaneously with dynamite by Johan Norrbin and Johan V. Ohlsson in 1867. Nevertheless, it was only after Texas City that the mining industry began to seriously consider this "new" explosive.

Soon after the advent of porous ammonium nitrate prills, introduced in the early 1950s, investigators realized that these prills readily absorbed just the right amount of fuel oil to produce an oxygen-balanced mixture that was both an inexpensive and an effective blasting agent, as well as being safe and simple to manufacture. This technology was widely adopted and soon constituted 85 percent of the industrial explosives produced in the United States.[23] With ANFO's cost and safety characteristics, it became practical for surface miners to drill larger boreholes and to utilize bulk ANFO delivery systems. Trucks were developed that could mix the AN and oil right on the mine site as needed and auger the ANFO directly into the boreholes. This technology has made possible the large open-pit copper mines and coal stripping operations seen around the world.

Nevertheless ANFO had two significant limitations: ammonium nitrate is very water-soluble, so wet boreholes readily deactivated the explosive; and ANFO's low density of 0.85 g/cc limited its bulk explosive strength. Many attempts were made to waterproof the product and add densifying agents, but with only limited success. However, because of this effort the next advance in industrial explosives technology was made by the invention of slurry explosives.

In the spirit of "If you can't beat 'em, join 'em", Dr. Melvin A. Cook[24] hit upon the idea of dissolving the ammonium nitrate in a small amount of hot water, mixing in fuels such as aluminum powder, sulfur, or charcoal, and adding a thickening agent to gel the mixture and hold the slurried ingredients in place. As this mixture cooled down, the AN salt crystals would precipitate, but the gel would preserve the close contact between oxidizer and fuels, resulting in a detonable explosive. Other oxidizers also could be added, and the density could be adjusted with chemical foaming agents to vary the bulk explosive strength of the product. With the addition of a cross-linking agent, the slurry or water gel could be converted to a semisolid material having significant water resistance. A picture of an early packaged gel is reproduced in Fig. 30.5.

The first commercial slurry explosives in 1957 required explosive sensitizers such as particulate TNT or smokeless powder to be detonable; but later with proper density control, better thickeners, and finely divided fuels, no self-explosive materials were required. Bulk slurry manufacturing trucks were developed where all of the ingredients were carried in separate compartments on the truck and then mixed on site and pumped directly into the boreholes. Figure 30.6 shows a typical pumping operation. Further formulation refinements in the late 1960s and early 1970s resulted in the development of slurry and water gel explosives for small-diameter, packaged applications where dynamite dominated. These formulations required mild

Fig. 30.5. The first packaged aluminized slurry explosive.

Fig. 30.6. A site-mixed slurry pumping operation at a coal mine.

sensitization with additives such as paint-grade aluminum powder or soluble energetic materials such as methylammonium nitrate (MAN), hydroxyethylammonium nitrate (MEAN), and ethyleneglycol mononitrate (EGMN) to make them cap-sensitive. Even though slurries were more expensive than ANFO, their flexibility and wide versatility made them applicable for all areas of mining from soft coal to hard iron ore and in all climatic conditions, even in water-filled holes. Thus, slurries provided a good adjunct to ANFO for areas with more stringent blasting conditions.

The latest significant development in industrial explosives actually was invented only a few years after slurries,[25] but full commercialization has occurred only within about the last fifteen years. Water-in-oil emulsion explosives involve essentially the same ingredients that slurry composite explosives do, but in a different physical form. Aqueous solutions of oxidizer salts, primarily ammonium nitrate, are emulsified into a relatively small amount of oil and emulsifier. Because of oxygen-balance constraints the volume ratio of the two liquid phases is about 90:10 for the aqueous and oil phases, respectively. This results in a high-internal-phase-ratio emulsion that has inherent viscosity without added thickeners due to the close packing of the oxidizer solution droplets. As each droplet is surrounded by a very thin film of organic fuel, this is a nearly ideal physical arrangement for efficient explosive reaction. Also, the very small oxidizer solution droplets (2–5-micron diameters are typical) tend to retard crystal growth, and the droplets supercool without crystallizing, thus preserving the oxidizer–fuel intimacy. Figure 30.7 is a photomicrograph of an emulsion explosive at a magnification of 400, showing the close packing of the oil-covered droplets. Because of the increased intimacy, emulsions generally have higher detonation velocities than slurries and are more sensitive to initiation; so they do not require chemical sensitizing agents for small-diameter applications. In addition, the continuous film provides emulsions with excellent water resistance. A bulk emulsion application under extreme water conditions at a quarry in Southeast Florida is shown in Fig. 30.8.

Depending upon the type of oil or wax used for the fuel phase, emulsions can be made with a thin, pumpable or pourable rheology for

Fig. 30.7. Photomicrograph of a repumpable emulsion explosive.

bulk applications, or can have a soft to firmly pliable consistency for packaged applications. With their unique physical and chemical properties, emulsions have replaced must slurry explosives. In recent years combinations of ANFO and emulsions have become popular, which combine the best features of both systems with mixtures ranging from pumpable to augerable, depending on the ratio of the two explosives.

The main developments in military types of explosives since World War II have been trends toward the use of plastic bonded explosives (PBXs) and the development of insensitive high explosives. Driving these trends are desires for increased safety and improved economics in the process of replacing aging TNT-based munitions and bomb fills. PBX explosives involve the coating of fine particles of molecular explosives such as RDX and HMX (1,3,5,7-tetranitro-1,3,5,7-tetrazacyclooctane) with polymeric binders and then pressing the resultant powder under vacuum to give a solid mass with the desired density. The final form or shape usually is obtained by machining. Explosives such as TATB (triaminotrinitrobenzene), nitroguanidine,[26] and HNS (hexanitrostilbene)[27] are of interest because of their high levels of shock insensitivity and thermal stability. The synthesis of many new, potentially explosive compounds is a very active and ongoing area of research,[28–29] but recently interest also has focused on composite explosives similar to those used by industry. Examples are EAK, a eutectic mixture of ethylenediamine dinitrate, ammonium nitrate, and potassium nitrate,[30] and nonaqueous hardened or cast emulsion-based mixtures.[31]

Fig. 30.8. A site-mixed emulsion pumping operation under very wet conditions.

TABLE 30.3 Some Properties of Common Explosives

Common Name	*Symbol*	*Composition*	*Molecular Weight*	*Density (g/cc)*	*Detonation Velocity (km/sec)*	*Detonation Pressure (kilobars)*	*Explosive Energy (cal/g)*
Primary Explosives							
Mercury fulminate		$Hg(CNO)_2$	284.7	3.6	4.7	220	428
Leadazide		$Pb(N_3)_2$	291.3	4.0	5.1	250	366
Silver azide		AgN_3	149.9	5.1	6.8		452
Lead styphnate		$C_6H(NO_2)_3O_2Pb$	468.3	2.5	4.8	150	368
Mannitol hexanitrate (Nitromannite)	MHN	$C_6H_8(ONO_2)_6$	452.2	1.7	8.3	300	1420
Diazodinitrophenol	DDNP	$C_6H_2N_4O_5$	210.1	1.5	6.6	160	820
Tetrazene		$C_2H_8N_{10}O$	188.2	1.5			658
Secondary Explosives							
Nitroglycerine	NG	$C_3H_5(ONO_2)_3$	227.1	1.6	7.6	253	1480
Pentaerythritol tetranitrate	PETN	$C(CH_2ONO_2)_4$	316.2	1.6	7.9	300	1510
Trinitrotoluene	TNT	$CH_3C_6H_2(NO_2)_3$	227.0	1.6	6.9	190	900
Ethyleneglycol dinitrate	EGDN	$C_2H_4(ONO_2)_2$	152.1	1.5	7.4		1430
Cyclotrimethylenetrinitramine (Hexogen or Cyclonite)	RDX	$C_3H_6N_3(NO_2)_3$	222.1	1.6	8.0	347	1320
Cyclotetramethylenetetranitramine (Octogen)	HMX	$C_4H_8N_4(NO_2)_4$	296.2	1.9	9.1	393	1350
Trinitrophenylmethylnitramine (Tetryl)		$(NO_2)_3C_6H_2N(CH_3)NO_2$	287.2	1.4	7.6	251	950
Nitroguanidine	NQ	$CH_4N_3NO_2$	104.1	1.6	7.6	256	721
Nitromethane	NM	CH_3NO_2	61.0	1.1	6.2	125	1188
Nitrocellulose	NC	Variable	—	1.4	6.4	210	950
Triaminotrinitrobenzene	TATB	$C_6H_6N_3(NO_2)_3$	258.2	1.8	7.9	315	829
Diaminotrinitrobenzene	DATB	$C_6H_5N_2(NO_2)_3$	243.2	1.6	7.5	259	993
Ethylenediamine dinitrate	EDDN	$C_2H_{10}N_4O_6$	186.1	1.5	6.8		948

Ethylenedinitramine (Haleite)	EDNA	$C_2H_6N_2(NO_2)_2$	150.1	1.5	7.6	266	1080
Picric acid		$C_6H_3O(NO_2)_3$	229.1	1.7	7.4	265	1000
Ammonium picrate (Explosive D)		$C_6H_6NO(NO_2)_3$	246.1	1.6	6.9		800
Picramide		$C_6H_4N(NO_2)_3$	228.1	1.7	7.3		1070
Hexanitrostilbene	HNS	$[C_6H_3C(NO_2)_3]_2$	450.2	1.7	7.1	200	1005
TACOT-Z		$C_{12}H_4N_8O_8$	388.2	1.6	7.2	181	980
Azobishexanitrobiphenyl	ABH	$C_{24}H_6N_{14}O_{24}$	874.4	1.8	7.6		
Dinitrotoluene	DNT	$CH_3C_6H_3(NO_2)_2$	182.1	1.5	5.0		700
Composition B		49/50/1 TNT/RDX/wax		1.7	8.0	294	1100
Pentolite		50/50 TNT/PETN		1.6	7.7	245	1100
Amatol		50/50 TNT/AN		1.6	6.5		950
Dynamite		Variable NG and various oxidizers and fuels		0.8–1.6	1.8–7.6	30–160	675–1090
Prilled AN—Fuel Oil	ANFO	94/6 AN/FO		0.8–0.9	1.5–4.0 (depends on diameter)		880
Slurries or water gels		Variable mixtures of oxidizers, fuels, and water		0.9–1.4	3.5–5.0		600–1200
Emulsions		Variable solutions of oxidizers in water and fuels		0.9–1.4	4.5–6.0		700–1100
Heavy ANFO		50–75% AN with 50–25% emulsion		1.1–1.3	4.0–4.5		755–815
Tertiary Explosives							
Mononitrotoluene	MNT	$CH_3C_6H_4NO_2$	137.1	1.2			
Ammonium perchlorate	AP	NH_4ClO_4	117.5	1.9	3.4	187	488
Ammonium nitrate	AN	NH_4NO_3	80.1	1.4	3.2		346

CLASSIFICATION OF EXPLOSIVES

The original classification of explosives separated them into two very general types: low and high, referring to the relative speeds of their chemical reactions and the relative pressures produced by these reactions. A low explosive is one that deflagrates or burns very rapidly instead of detonating. Low explosives generally generate pressures below 50,000 psi, whereas high explosives detonate and produce pressures of up to 5 million psi. The stable detonation rate of a high explosive has been characterized as being faster than the speed of sound in the explosive, and the deflagration of a low explosive is less than the speed of sound. This classification still is used but is of limited utility because the only low explosives of any significance are black powder and smokeless powder. All other commercial and military explosives are high explosives.

High explosives are classified further according to their sensitivity level or ease of initiation. Actually sensitivity is more of a continuum than a series of discrete levels, but it is convenient to speak of primary, secondary, and tertiary high explosives. Primary explosives are the most sensitive, being readily initiated by heat, friction, impact, or spark. They are used only in very small quantities and usually as an initiator as part of an explosive train involving less sensitive materials, such as in a blasting cap. They are very dangerous materials to handle and must be manufactured with the utmost care, generally involving only remotely controlled operations. Mercury fulminate, used in Nobel's first blasting cap, is in this category, as is the more commonly used lead azide. On the other end of the spectrum are the tertiary explosives that are so insensitive that they generally are not considered explosive. In this category are the oxidizers ammonium nitrate and ammonium perchlorate, which are quite safely manufactured as commodity chemicals. The former is in daily use by consumers as a common fertilizer for lawns and gardens. The latter is used as a component in rocket motors, but its explosive potential recently was demonstrated in a spectular accidental detonation.[32]

By far the largest grouping is secondary explosives, which includes all of the major military and industrial explosives. They are much less easily brought to detonation than primary explosives and are less hazardous to manufacture. Beyond that, however, generalizations are difficult because their sensitivity to initiation covers a very wide range. Generally the military products tend to be more sensitive and the industrial products less sensitive, but all are potentially hazardous and should be handled and stored as prescribed by law. Table 30.3 lists some of the more prominent explosives of each type, along with a few of their properties.

For industrial applications, secondary explosives are subdivided according to their initiation sensitivity into two classes: Class A and blasting agents. Class A explosives are sensitive to initiation by a blasting cap and usually are used in relatively small-diameter applications of 1- to 3-inch boreholes. Blasting agents are high explosives that are not initiated by a Standard #8 electric blasting cap under test conditions defined by the U.S. Department of Transportation, and that pass other defined tests designed to show that the explosive is "so insensitive that there is very little probability of accidental initiation to explosion or of the transition from deflagration to detonation."[33] Being less sensitive, blasting agents are generally used in medium- and large-diameter boreholes and in bulk applications. Dynamites are always Class A, but other composite explosives made from mixtures of oxidizers and fuels can be made either Class A or blasting agent, depending upon the formulation and the density. Density plays a significant role in the performance of most explosives, and this is especially true for slurry and emulsion explosives where the density may be adjusted by air incorporation, foaming agents, or physical bulking agents, irrespective of the formulation. Blasting agent classification is of interest because regulations governing transportation, use, and storage are less stringent for blasting agents than for Class A explosives. (Propellants and fireworks are

classified by the DOT as Class B Explosives, and blasting caps and detonating cord as Class C.)

STRUCTURAL CHARACTERISTICS OF EXPLOSIVES

The number of potentially explosive compounds is virtually unlimited. A recent listing by the U.S. Bureau of Alcohol, Tobacco and Firearms of explosive materials under federal regulation[34] numbered 225, and many of the items listed were broad, general categories. The ten-volume *Encyclopedia of Explosives and Related Items* compiled by the U.S. Army Picatinny Arsenal over a 25-year period contains several thousand entries. New organic molecular explosives are being synthesized continually; composite explosives, such as current commercial products that are mixtures of oxidizers and fuels, present an infinite number of possible combinations. The complexity of trying to comprehensively list the chemical structures of explosives is shown by a 1977 reference that listed 13 separate categories just for primary explosives.[35] However, the majority of the most important explosives can be grouped into a few classes sharing common structural features that are of value to researchers in understanding and predicting explosive properties.

The following seven categories,[36] updated to include the relatively recent fluoroderivatives,[28] appear to be the most encompassing. Many explosives may contain more than one category, but not every compound that contains one of these groups is necessarily an explosive.

1. —N=O
2. —N—N—, —N=N—, and —N≡N
3. —C≡N— and —C≡N
4. —C≡C—
5. —Cl=O
6. —N—X, where X = Cl, F, I
7. —O—O—

Category 1 is by far the largest. It includes nitro groups, both aliphatic and aromatic; nitrate esters; nitrate salts; nitramines; and nitrosamines. Nearly all of the explosives listed in Table 30.3 fall into this category. Prominent examples are: nitromethane, an aliphatic nitro compound; TNT, an aromatic nitro compound; nitroglycerin and PETN, nitrate esters; EDDN and ammonium nitrate, nitrate salts; and RDX and HMX, nitramines. Category 2 represents the hydrazine, azo, diazo, and azide compounds, both organic and inorganic. Hydrazine, tetrazene, and lead azide are examples of this group. Category 3 is represented by the explosives mercury fulminate and cyanogen, respectively. Acetylene and metallic acetylide salts constitute Category 4. Category 5 consists mainly of inorganic and organic ammonium salts of chloric and perchloric acid, but would also include various chlorine oxides. Category 6 is generalized to include most of the amine halogens, nitrogen triiodide being a classic example. Also, considerable new synthetic work has focused on inserting the energetic difluoroamine groups into various organic molecules to form explosives that fall into this category. Category 7 includes organic peroxides and ozonides as well as hydrogen peroxide itself.

Commercial industrial explosives such as dynamites, slurries, and emulsions are included in these categories because their major components, nitrate esters and nitrate and perchlorate salts, are listed. However, mixtures of fuels and oxygen or other gases that may be explosive at certain ratios are not covered, including the liquid oxygen explosives that saw limited application earlier in the twentieth century.

EXPLOSIVES MANUFACTURING

Details of the synthesis and larger-scale production of a number of molecular explosives including dynamites are given in the four-volume series by Urbanski (*Chemistry and Technology of Explosives*, Pergamon Press, 1964–84) and in various military books such as reference 8. Formulations of commercial slurries and emulsions generally are considered proprietary and are described mainly in the patent literature. Some specific

examples of prominent explosives with general preparation methods are given below.

TNT (2,4,6-Trinitrotoluene)

TNT is no longer manufactured commercially in the United States, but is manufactured in significant quantities at several government plants because it is still an important military explosive. It is produced commercially in Canada and other countries and is imported into the United States for use in cast boosters to initiate industrial blasting agents. In a relatively straightforward process, TNT is made by the direct trinitration of toluene with nitric acid. Most modern processes are set up for continuous production in a series of nitrators and separators with the nitrating acid flowing countercurrently. This procedure avoids having to isolate the intermediate mono- and dinitration products and may also employ continuous purification and crystallization, being carried out simultaneously with production.

Mixed nitric and sulfuric acids sometimes are used with the addition of SO_3 or oleum. The sulfuric acid or oleum helps drive the reaction to completion by removing the water produced by nitration and by dehydrating nitric acid to form the more reactive nitronium ion (NO_2^+). Because toluene is not very soluble in the acid, powerful agitation is required. The spent acid is removed in successive separation steps, and the sulfuric acid is reused after the addition of more nitric acid. The molten TNT product is purified with multiple water and sodium sulfate washes, which produce significant quantities of "yellow water" and "red water" waste streams, respectively, that must be properly handled to avoid environmental problems. The low melting point of TNT (80–82 °C) is ideal for melt casting, and TNT usually is employed as a mixture with other higher-melting explosives such as PETN, RDX, HMX, and tetryl. This feature and the excellent chemical stability of TNT have made it the most popular and widely used military explosive in the world.

RDX and HMX

Both RDX and HMX are cyclic nitramines made by nitrolysis of hexamethylene tetramine (HMT). Their good thermal stabilities, high melting points (>200 °C), and high energy properties make these crystalline compounds popular as projectile and bomb fills and for use in cast boosters and flexible, sheet explosives. HMX has superior detonation properties and a higher melting point than RDX, but it is more difficult and more expensive to manufacture. Reaction 1 shows the formation of RDX by the action of nitric acid on HMT. Schematically, RDX formation can be pictured as nitration of the three "outside" nitrogen atoms of HMT (in more accurate, three-dimensional representations all four nitrogens are equivalent) with removal of the "inside" nitrogen and methylene (—CH_2—) groups. Ammonium nitrate (NH_4NO_3) and formaldehyde (CH_2O) are produced as by-products but can be used to form more RDX with the addition of acetic anhydride, as shown in Reaction 2. In actual practice these two reactions are run simultaneously, as shown in the combined reaction to produce approximately two moles of RDX for each mole of HMT.

HMX was discovered as an impurity produced in the RDX reaction. It is composed of an eight-membered ring rather than the six-membered ring of RDX. The latter is more readily formed than the eight-membered ring, but with adjustment of reaction conditions (lower temperature and different ingredient ratios), HMX formation can be favored. Schematicaly its formation can be pictured by nitration of all four nitrogens in hexamethylene tetramine and removal of two methylene groups as indicated in Reaction 3. To obtain pure HMX the RDX "impurity" must be removed by alkaline hydrolysis or by differential solubility in acetone.

HNS (2,2′,4,4′,6,6′-Hexanitrostilbene)

This is a relatively new explosive having been prepared unequivocally for the first time in the early 1960s.[37] It is of interest primarily

Reaction 1

Reaction 2

$$3CH_2O + 3NH_4NO_3 + 6(CH_3CO)_2O \rightleftharpoons (CH_2{\cdot}N{\cdot}NO_2)_3 + 12CH_3COOH$$

Formaldehyde, Ammonium nitrate, Acetic anhydride, RDX, Acetic acid

Combined Reaction

$$(CH_2)_6N_4 + 4HNO_3 + 2NH_4NO_3 + 6(CH_3CO)_2 \rightleftharpoons 2(CH_2{\cdot}N{\cdot}NO_2)_3 + 12CH_3COOH$$

Reaction 3

for two reasons: (1) its high melting point (316 °C) and excellent thermal stability, and (2) its unique crystal-habit-modifying effects on cast TNT. The former makes HNS useful in certain military and space applications as well as in hot, very deep wells, and the second property is used to improve TNT castings. It can be manufactured continuously by oxidative coupling of TNT as shown below.

This relatively simple process from readily available TNT and household bleach (5% NaOCl solution) has been shown to involve a series of intermediate steps that give HNS in only low to moderate yields (30–45%) with many by-products. Although it also involves the use of expensive organic solvents that must be recovered, this synthesis is used commercially.[38,39] Studies to improve this process constitute an active area of research.

TATB (1,3,5-Triamino-2,4,6-trinitrobenzene)

This highly symmetrical explosive molecule has even higher thermal stability than HNS (greater than 400 °C) and has become of special interest in the last two decades because of its extreme insensitivity.[40,41] Because its accidential initiation is highly unlikely, TATB has been used in nuclear warheads and is being explored in plastic bonded systems for a number of military and space applications.[26] Currently it is manufactured in large-scale batch processes that are little change from its original synthesis over 100 years ago. The two-step process involves trinitration of trichlorobenzene followed by amination to displace the chlorine groups as shown below.

Trichlorobenzene

Trinitrotrichlorobenzene TATB

Both steps require high temperature and considerable reaction time but give 80 to 90 percent yields. The major problem areas are chloride impurities in the final product and the excessively fine particle size of the final product. Because TATB is highly insoluble in most solvents, it is difficult to purify the product or to change its particle size by recrystallization. Also the starting material is expensive and not very readily available.

Recently a similar synthetic procedure starting with 3,5-dichloranisole was reported.[42]

DDNP (2-Diazo-4,6-dinitrophenol)

This yellow-to-brown crystalline material (melting point 188 °C) is a primary explosive used as the initiator charge in electric blasting caps as an alternative to lead azide. It is less stable than lead azide but much more stable than lead styphnate, and is a stronger explosive than either of them because it does not contain any metal atoms. DDNP is also characterized as not being subject to dead pressing (tested at pressures as high as 130,000 psi). It was the first diazo compound discovered (1858) and was commercially prepared in 1928. It is manufactured in a single-step, batch process by diazotizing a slurry of sodium picramate in water.

Sodium Picramate

DDNP

The structure shown is convenient for visualization purposes, but DDNP actually exists in several tautomeric forms with form (2) apparently predominating.

(3) (1) (2)

The sodium picramate starting material is itself explosive, but is commercially available as a chemical intermediate. It can be made by the reduction of picric acid with reducing agents such as sodium sulfide. The key to making useful DDNP is to control the rate of diazotization so that relatively large, rounded crystals are formed instead of needles or platelets that do not flow or pack well.

PETN (Pentaerythritol Tetranitrate)

Although known as an explosive since 1894, PETN was used very little until after World War I when the ingredients to make the starting material became commercially available. The symmetrical, solid alcohol starting material, pentaerythritol, is made from acetaldehyde and formaldehyde, which react by aldol condensation under basic catalysis followed by a crossed Cannizzaro disproportionation to produce the alcohol and format salt. Although the reaction takes place in a single mixture, it is shown below in two steps for clarity.

For PETN manufacture the pentaerythritol starting material can be readily purchased as a commodity chemical from commercial suppliers. The nitration is relatively simple, involving only nitric acid (96–98%) and the solid alcohol added slowly with mixing and cooling. PETN is not very soluble in nitric acid or water and is readily filtered directly from the acid or after dilution of the acid with water. Water washing and recrystallization from acetone–water mixtures give the desired particle size ranges and the desired purity. PETN can be made either batchwise or continuously for large-scale production.

Pure PETN is a white, crystalline solid with a melting point of 141.3 °C. Because of its symmetry it is said to have higher chemical stability than all other nitrate esters.[43] Relatively insensitive to friction or spark initiation, PETN is easily initiated by an explosive shock and has been described overall as one of the most sensitive, non-initiating, military explosives.[44] As with most explosives, the detonation velocity of PETN varies with the bulk density of the explosive. Figure 30.9 is a plot of measured PETN velocities at different pressed powder and crystalline densities, which illustrates the wide range of detonation rates achievable. Most military applications of PETN have been converted to RDX because of its greater thermal stability. However, in industry PETN is widely used as the major component in cast boosters for initiating blasting agents, as the explosive core in detonating cord, and as the base load in detonators and blasting caps. For safety in handling, PETN is shipped in cloth bags immersed in water/alcohol mixtures and dried just before use.

NG (Nitroglycerin or Glyercol Trinitrate)

This nitrate ester is one of only a very few liquid molecular explosives that are manufactured commercially. It is a clear, oily liquid that freezes when pure at 13 °C. As seen in the historical section, the first practical use of NG was in dynamites, where it is still used today more than 100 years later. It also is used as a component in multibased propellants and as a medicine to treat certain coronary ailments. This latter usage is attributed to NG's ability to be rapidly absorbed by skin

Fig. 30.9. PETN detonation velocity vs. density.

contact or inhalation into the blood, where it acts as a vasodilator. (At high exposure levels such as occur in dynamite manufacture and handling, this property is responsible for the infamous powder headache.) NG is undoubtedly the most sensitive explosive manufactured in relatively large quantities. Its sensitivity to initiation by shock, friction, and impact is very close to that of primary explosives, and extreme safety precautions are taken during manufacture. Pure glycerin is nitrated in very concentrated nitric and sulfuric acid mixtures (typically a 40/60 ratio), separated from excess acid, and washed with water, sodium carbonate solution, and water again until free from traces of acid or base. Pure NG is stable below 50 °C, but storage is not recommended. It is transported over short distances only as an emulsion in water or dissolved in an organic solvent such as acetone. Traditionally it has been made in large batch processes, but safety improvements have led to the use of several types of continuous nitrators that minimize the reaction times and quantities of explosive involved. Because of its sensitivity, NG is utilized only when desensitized with other liquids or absorbent solids or compounded with nitrocellulose.

Dynamite

Dynamite is not a single molecular compound but a mixture of explosive and nonexplosive materials formulated in cylindrical paper or cardboard cartridges for a number of different blasting applications. Originally Nobel simply absorbed NG into kieselguhr, an inert diatomaceous material, but later he replaced that with active ingredients—finely divided fuels and oxidizers called dopes. Thus, energy is derived not only from the NG, but also from the reaction of oxidizers such as sodium nitrate with the combustibles. Its manufacture

involves mixing carefully weighed proportions of NG and various dopes to the desired consistency and then loading preformed paper shells through automatic or semiautomatic equipment. Because dynamites represent the most sensitive commercial products produced today, stringent safety precautions, such as the use of nonsparking and very-little-metal equipment, good housekeeping practices, limited personnel exposure, and barricaded separations between processing stations, are necessary.

Table 30.4 lists the common general types of dynamites with their distinguishing features. The straight dynamites and gelatins largely have been replaced by the ammonia dynamites and ammonia gelatins for better economy and safety characteristics. The gelatins generally have higher detonation velocities and better water resistance than other types but are more costly and are not suitable for breaking soft materials. Permissible explosives are those tested by the U.S. Bureau of Mines and approved by MSHA for use in underground coal mines. They are designed to reduce the probability of accidental initiation of combustible mine gases or dusts. The fuel dopes used in dynamites have changed very little over the years, and are composed mostly of wood pulps, nut meals, sawdust, flour, starch, and cereal by-products. Early safety problems with "frozen" dynamites were overcome in 1925 by the addition of ethyleneglycol dinitrate (EGDN) to nitroglycerin to lower its freezing point. Today the "NG" used in dynamites is actually a mixture of EGDN and nitroglycerin (formed by nitrating mixtures of the two alcohols), in which nitroglycerin is usually the minor component because EGDN is cheaper than NG and has better heat stability. The NG content of the various dynamite formulations ranges from as little as 5 percent to as much as 90 percent of the composition.

In 1980 a new commercial line of "non-NG" dynamites was introduced that employs higher molecular weight nitrate esters to reduce the headache potential of standard dynamites.[45] Mixtures of metriol trinitrate (MTN) and diethyleneglycol dinitrate (DEGDN), though less sensitive, perform similarly to the NG mixtures in explosive formulations. Chemical structures for these compounds are shown below, along with that of EGDN for comparison.

TABLE 30.4 General Types of Dynamite

1. Straight dynamite	Granular texture with NG as the major source of energy
2. Ammonia dynamite ("extra" dynamite)	Ammonium nitrate replacing part of the NG and sodium nitrate of the straight dynamite.
3. Straight gelatin dynamite	Small amount of nitrocellulose added to produce soft to tough rubbery gel.
4. Ammonia gelatin dynamite ("extra" gelatin)	Ammonium nitrate replacing part of the NG and sodium nitrate of the straight gelatin.
5. Semigelatin dynamite	Combination of types 2 and 4 with in-between properties.
6. Permissible dynamite	ammonia dynamite or gelatin with added flame retardant.

Slurry and Emulsion Explosives

Water-based commercial explosives that are mixtures of oxidizers and fuels represent a tremendous advance in economy, safety, and versatility. The components can be varied to change detonation properties and energy levels to suit the application, as can dynamites, but they do not contain NG. Water-based explosives have been designed that cover a very broad range of applications from small-diameter (less than 1 inch) to large-diameter (greater than 6 inches); from packaged to bulk; from cap-sensitive to booster-sensitive; from pourable liquids to rigid, hard solids.

Although they are made from basically the same ingredients, water-based explosives fall into two distinct types: slurries or water gels, and emulsions.[46]

Slurries are made by dissolving part or all of the oxidizer salts in a certain proportion of water, thickening this solution or slurry with a hydrating polymer thickener, and mixing in the soluble and/or insoluble fuels. Usually a cross-linking agent is added to firm up the system and hold any solids in place, and the slurry is pumped into a package or directly into a borehole. Some common slurry ingredients are listed in Table 30.5. Soluble sensitizers are used by some manufacturers to enhance slurry detonation characteristics, particularly in small diameters. In their anhydrous state, some of these sensitizers may be explosive by themselves, but usually they are formed *in situ* or added as aqueous solutions. Generally, slurries are made at temperatures of 40 to 60 °C to minimize the amount of water required to dissolve the oxidizer salts. When these slurries cool, the salts precipitate, producing the characteristic rheology or body of these thickened gels.

Emulsion explosives are made by completely dissolving the oxidizer salts in water at 60 to 90 °C or higher temperatures and adding the solution with agitation to a liquid fuel consisting of a water-in-oil emulsifying agent dissolved in oil or melted wax. Other soluble and insoluble fuels may be added, but thickeners and soluble sensitizers generally are not needed. The fuel phase of this two-phase system may be heated to a similar temperature range in order to employ solid paraffin or microcrystalline waxes. Because the tiny oxidizer solution droplets in these emulsions tend to remain fluid upon cooling, the rheology of the final product is mostly influenced by the viscosity of the oil–fuel phase. If very thin oils are used, pumpable or even pourable products result. If very thick oils or waxes are used, tacky to firmly rigid products are obtained.

Slurries and emulsions may be conveniently manufactured by either batch or continuous processes, and their inherent safety lessens the need for the more stringent safety measures taken in making molecular explosives. Because these composite explosives can be made entirely from what are considered nonexplosive ingredients, mobile manufacturing units for on-site production have been developed. Figure 30.10 is a schematic of a truck unit designed to hold the various liquid and solid ingredients, which are augered or pumped to a mixer and then pumped through a long hose to the bottom of the borehole. This method allows the formulation or energy level to be changed during loading and also allows any unused ingredients to be returned. At their natural densities of up to 1.5 g/cc, these

TABLE 30.5 Typical Slurry Ingredients

Oxidizers	*Soluble Fuels*
ammonium nitrate (AN)	ethylene glycol
sodium nitrate (SN)	methyl alcohol
calcium nitrate	sugar
sodium perchlorate	alkyl amines
Thickeners	*Insoluble Fuels*
guar gum	particulate aluminum
starch	coal dust
synthetic water-soluble polymers	charcoal
	sulfur powder
Soluble Sensitizers	fuel oil
ethyleneglycol mononitrate (EGMN)	
monoethanolammonium nitrate (MEAN)	
monomethylammonium nitrate (MAN)	
ethylenediammonium dinitrate (EDDN)	

Fig. 30.10. Slurry or emulsion site-mixed pump truck schematic.

water-based explosives are generally too dense to detonate effectively in the diameters of interest. Therefore their densities are lowered to the range shown in Table 30.3 through the use of chemical foaming agents or low density fillers, such as expanded perlite particles or tiny glass microballoons (10–200 microns in diameter). Products with various energy levels are produced by changing the amount of water in the formulations and by adding varying amounts of granular or powdered aluminum as a high energy fuel.

Some generalized formulas for slurry and emulsion explosives are shown in Table 30.6.

TABLE 30.6 Typical Slurry and Emulsion Formulations

	Slurries		*Emulsions*		
AN	57.3	53.5	60.0	70.7	78.4
SN	9.3	8.6	18.1	10.7	—
MAN	—	21.3	—	—	—
Ethylene glycol	1.9	—	—	—	—
H_2O	18.9	8.1	11.8	7.3	13.5
Thickener	0.8	1.8	—	—	—
Emulsifier	—	—	1.0	0.8	1.5
Oil/wax	—	—	5.2	2.2	5.5
Aluminum	10.2	2.5	—	5.7	—
Sulfur	—	2.0	—	—	—
Density agent	1.6	2.2	3.9	2.6	1.1

CONSUMPTION OF COMMERCIAL EXPLOSIVES

Data on military explosive production are not readily available, but the Bureau of Mines has collected data annually from industrial companies since 1922 and published them as listed in reference 3. These values, combined with totals from reference 20 for black powder and liquid oxygen blasting agents, are plotted in Fig. 30.11, which shows that, except for the depression years, the apparent consumption of explosives in the United States for commercial use was fairly constant for the first 35 years, at between 400 and 500 million lb/year. Then consumption doubled over the next ten years, from 1946 to 1956, and with the advent of ANFO, doubled again in each of the next 10- to 12-year periods, reaching 4 billion lb by 1979. Since then there has been some volatility due to changing market factors, but consumption in general has continued to increase, and in 1989 was 4.8 billion lb. Over 83 percent of this total was ANFO explosives. The various types of dynamite amount to less than 3 percent of the total, with the newer slurries and emulsions comprising more than 13 percent.

Of the industries that consume explosives, coal mining is the largest consumer, accounting for 66 percent of all the commercial explosives

Fig. 30.11. The consumption of explosives in the United States, 1912–89.

Fig. 30.12. Explosive usage by industry, 1980–89 (million lb).

produced in 1989. Metal mining used 10 percent, quarrying and nonmetallic mining utilized about 13 percent, and construction and other uses accounted for the remaining 11 percent. Figure 30.12 gives the industry distribution for 1980 through 1989. These data show that explosives continue to play a very important and vital role in modern society and are undoubtably among the most efficient uses of energy known.

Fig. 30.13. 1987 OSHA reportable incidents per 100 workers.

The emphasis of the explosives industry on safety in the manufacturing and handling of explosives has led to its being one of the safest industrial environments. A comparison of 1987 work injury rates for the explosives industry, for all industries, and for several selected industries is shown in Fig. 30.13. The data listed are OSHA recordable occupational injuries and illnesses per 100 employees as reported by the National Safety Council.[47] The development of modern blasting agents and explosives, such as ANFO, slurries, and emulsions, that contain no self-explosive ingredients has further advanced safety of application, and has lowered costs and increased the options available to the consumer. The current level of R&D effort is likely to continue this trend and also to lead to extension of this technology into military applications.

ROCKET PROPELLANTS

A rocket is a device that uses the expulsion of internally generated gases as a source of motive power. The gases used for propelling the rocket are generated by chemical reaction of a fuel and an oxidizer. The force that acts against a rocket as gases are expelled is called the thrust. Because rockets carry their own fuel and oxidizer and do not rely on air, the thrust from reaction (combustion) of the propellant chemicals will act in a vacuum. Thus, rockets, unlike internal combustion engines, are capable of providing power in space as well as in the earth's atmosphere.

The use of rockets has been traced to thirteenth century China, but it was not until the development of the liquid-fueled V-2 in Germany during World War II that a practical long-range missile using rocket propulsion was achieved. Work during the early twentieth century by such pioneers as Robert Goddard in the United States and Herman Oberth in Germany provided the basis for the successful German effort and the spectacular space exploration studies that followed. The launching into orbit of the *Sputnik* satellite from the Soviet Union in 1957 was the initial event in a huge expansion of rocket development efforts in recent years. These developments have resulted in rockets used for three principal applications:

- Space exploration and satellite launching
- Strategic missiles
- Tactical missiles

Space exploration efforts have been very well publicized in recent years, and have included such notable developments, in addition to Sputnik, as the launching of manned rockets (with the first astronauts Yuri Gagarin in the Soviet Union and Alan Shepard in the United Sates), the Apollo missions to the moon (with Neil Armstrong's momentous first step), the Russian and

American space station and Skylab efforts, the U.S. space shuttle program, and the exploration of the solar system by such spacecraft as the Russian Venera and the U.S. Pioneer, Mariner, and Voyager. Perhaps less well publicized, but of great commercial and strategic importance, has been the launching of satellites for purposes of communication, mapping, and surveillance. Launch vehicles for the U.S. space program have included the Atlas Agena, Delta, Juno, Saturn, Scout, Thor, and Titan rockets.

The space exploration efforts were paralleled in the United States and the Soviet Union by the development of rocket-powered missiles for strategic military use. Such U.S. systems as the Air Force Minuteman, Peacekeeper, and Small ICBM and the Navy submarine-launched Polaris and Trident are widely deployed.

The use of missiles for tactical military applications has also been an area of major development since World War II. Among the first such applications were the JATO (rocket assisted takeoff) units used to provide power to boost launching of airplanes. Tactical missiles have become an important component of weaponry and include U.S. rockets such as the Navy Sidewinder, the Army Hawk and Hellfire, and the Air Force Sparrow and Phoenix.

PRINCIPLES OF ROCKET PROPULSION

The flight of rockets is based on the thrust achieved by expelling gases from the aft end of the missile; this provides a forward impetus. A schematic diagram of a simple rocket is shown in Figure 30.14. Combustion of the propellant causes pressurization of the chamber by hot gases; the pressure from the gases is counterbalanced by the strength of the chamber. At a narrow opening, the throat, gases are allowed to escape, providing thrust. If there were no expansion cone, and gases were expelled at the throat, the force F, acting to propel the rocket would be:

$$F = A_t P_c$$

where A_t is the area of the throat and P_c is the chamber pressure. When an expansion cone is present, a new term called the thrust coefficient, C_f, enters the equation:

$$F = A_t P_c C_f$$

The value of C_f depends on the ratio of the chamber pressure to the pressure at the exit plane, and on the ratio of the area of the throat to the area at the exit plane. The optimum performance of a rocket results when the pressure at the exit plane, P_e, is equal to the pressure of the surrounding atmosphere (which is one atmosphere for

Fig. 30.14. Schematic drawing of a simple rocket.

firings at sea level and zero atmospheres in space).

Because thrust is dependent on motor design and the rate of propellant combustion, it is not a convenient measure of propellant effectiveness. A parameter that is used to compare effectiveness is the specific impulse, I_{sp}, which is equivalent to the force divided by the mass flow rate of the propellant:

$$I_{sp} = \frac{Ft}{w} = \int \frac{F}{\dot{w}}$$

where $\dot{w}$ is the weight flow rate of the propellant, w is the total weight of the propellant, and t is the time. Because the impulse is dependent on a variety of parameters, it is customary to use the standard specific impulse, I^0_{sps}, which is the value of the specific impulse for an ideal rocket motor fired at 1000 psi, exhausting to 14.7 psi, with no heat loss, and with a nozzle of 0° half-angle. Frequently, measured or delivered impulse, I_{spd}, values from motor firings will be converted to I^0_{spd} for comparison with previous firings and with expectations. The ratio of delivered to predicted impulse is termed the efficiency. In engineering units, specific impulse is given in pounds(force) seconds/pound(mass)$^{-1}$. A thorough yet succinct discussion of the physics and thermodynamics of rocket propulsion is found in Sutton.[48]

The prediction of rocket propellant specific impulse, as well as impulse under other conditions, may be reliably accomplished by calculation using as input the chemical composition, the heat of formation, and the density of the component propellant chemicals. Not only impulse but also the composition of exhaust species (and of species in the combustion chamber and the throat) may be calculated if the thermodynamic properties of the chemical species involved are known or can be estimated. The present standard computer code for such calculations is that described by Gordon and McBride.[49] Theoretical performance predictions using such programs are widely used to guide propellant formulation efforts and to predict rocket propellant performance; however, verification of actual performance is necessary.

TYPES OF PROPELLANTS

The two principal types of rocket propellant in general use are solid propellants and liquid propellants. Solid propellants are chemical compositions that burn on exposed surfaces to produce gases for rocket power. Liquid propellants rely on pumping or pressurized flow of stored liquids to the combustion chamber. The choice between solid and liquid propellants for a specific application depends on a variety of considerations; to date, many of the strategic missiles and most of the tactical missiles rely on solid propellants because of the lower cost of the rocket and the greater storability of the propellant. On the other hand, large space vehicles and rockets fired for maneuvering in space use liquid propellants, in part because of the ready controllability of liquid systems. Less widely used are hybrid rockets, which use solid fuel and liquid oxidizer. If a single chemical compound (for example, nitromethane) containing both oxidizing and reducing functions in the same molecule is employed to power a rocket, it is called a monopropellant. If two chemicals combine to provide the propulsion, they form a bipropellant system.

SOLID PROPELLANTS

A solid propellant rocket motor is quite simple in concept, although in practice a complete motor is more complex. As shown in Figure 30.15, the rocket propellant is contained within a case, which may be metal or a reinforced high-performance composite. Frequently, the case is internally shielded by a bonded layer of insulation. The insulation is coated with a liner that bonds the propellant to the insulation. The integrity of the propellant-to-liner bond is of utmost importance; failure at this interface during a motor firing can result in a sudden increase in the area of propellant surface exposed to

Fig. 30.15. A solid propellant rocket motor.

combustion, with potentially catastrophic results.

The bore or perforation of the propellant grain is a major factor in determining the ballistic performance of a rocket. In the simplest case, the grain has no perforation, and the burning is restricted to the end of the grain; the resulting end-burning rockets have a relatively long burning duration with low thrust. More commonly, a perforation extends through the grain (which is center-perforated) and may have a cylindrical, star, cross-shaped, or more complex profile. The configuration of the grain is used to control the burning behavior of the propellant; the more surface area there is exposed, the more rapidly propellant will be consumed. The length of time a rocket motor will burn is governed by the perforation geometry, the web thickness of the propellant (distance from perforation to liner), and the burning rate of the particular rocket propellant. The pressure–time curve resulting from motor burning may be neutral (a single pressure is achieved and maintained throughout the burn), progressive, or regressive. Progressive burning leads to acceleration, whereas regressive burning gives lower pressure as the firing progresses. Tactical missiles frequently are manufactured with two types of propellant: a rapid-burning boost propellant to provide initial acceleration and a slower-burning sustain propellant to complete the desired flight profile. The trajectory of missiles as a function of propellant properties, grain configuration, and missile design may be reliably predicted or simulated by sophisticated computer calculations.

Single and Double-base Propellants

Early solid rocket propellants were based on processing similar to that used in the rubber industry, with propellants extruded into the desired grain configuration. Propellants that have been successfully manufactured by this technique include nitrocellulose (single-base) and nitrocellulose–nitrato ester (double-base) materials. Double-base propellants contain both nitrocellulose and nitroglycerin (NG) as the principal components; additionally, chemicals such as stabilizers, plasticizers, and burning rate modifiers may be added as appropriate. Other nitrato esters also may be in double-base systems. Extruded propellants usually are limited to small grain diameters (<12 in.) by the size of the equipment required for extrusion, or by the difficulties of solvent removal if a solvent-based process is employed.

A processing advantage is achieved with the castable double-base systems; a rocket chamber is filled with particulate nitrocellulose (casting powder), which then is treated with NG or a mixture of nitrato esters. The nitrocellulose is swollen by the nitrato ester to give the final propellant, a tough material with relatively low elasticity.

Double-base propellants may be formulated to include fuels such as aluminum metal, oxidizers such as ammonium perchlorate (AP), or energetic materials such as the nitramines RDX or HMX. The resulting compositions are termed composite-modified double-base propellants. A further modification, the addition of a polymer that is curable with a low molecular weight curing agent, allows formulation of propellants with much improved mechanical properties over the temperature range of usage. Such propellants are termed elastomer-modified composite double-base (EMCDB) propellants.

Composite Propellants

In recent years, the great majority of solid rockets have utilized composite propellants. Most composite propellants are based on a solid oxidizer and a curable liquid polymeric binder. The binder also serves as fuel. Optionally, a metallic fuel also may be used. The propellant components are mixed together, and then the binder is cured to give the tough, flexible, elastomeric (rubbery) solid propellant required for modern missile use. An excellent source of information on the formulation of solid propellants is a report by Oberth.[50]

Oxidizer. The major component of most composite solid propellants is the oxidizer. By far the most important oxidizer used in recent years is ammonium perchlorate (AP). This chemical possesses the desirable properties of high density, good thermal stability and oxygen availability, and relatively low reactivity and cost. Properties of AP and several other materials that are used as oxidizers are summarized in Table 30.7.

Ammonium nitrate (AN) has been considered as an oxidizer for many applications; but its principal use to date is in gas generator propellants, where generation of gases to provide initial motor pressurization or to power turbines is required. Widespread use of AN has been hindered by a propensity of the compound to undergo crystalline phase transitions involving a large (4%) volume change of the oxidizer, with concomitant oxidizer particle degradation, upon warming or cooling. Phase-stabilized ammonium nitrates (PSANs), which avoid this difficulty, have been developed but for various reasons have not yet found widespread use. An area of potential application for AN is in insensitive minimum smoke propellants (see Fig. 30.16).

Sodium nitrate has recently become of interest as a co-oxidizer for scavenger propellants. These propellants contain sufficient alkali metal (sodium) in the formulation to react with the chloride generated during combustion of the major oxidizer, ammonium perchlorate. Sodium chloride, rather than hydrochloric acid, is produced in the exhaust stream as the final chlorine-containing reaction product. The benefit of scavenging the chloride ion is lowered exhaust toxicity; however, a substantial loss of propellant impulse results.

The energetic nitramines, RDX and HMX, provide excellent impulse and nonsmoky exhaust, but their use gives propellant that may detonate when subjected to shock. The possibility of unwanted violent burning or detonation of propellants during transport or

TABLE 30.7 Properties of Oxidizers used in Solid Propellants

Material	*Formula*	*Molecular Weight*	*Density, g/cm³*	*Heat of Formation ΔH_f° @298°K, cal/100g*	*g Atom Oxygen/100g*
AP	NH_4ClO_4	117.4	1.95	−60.21	3.404
AN	$N_2H_4O_3$	80.0	1.725	−109.12	3.748
NaN	$NaNO_3$	85.0	2.26	−131.23	3.530
KP	$KClO_4$	138.6	2.53	−74.49	2.887
RDX	$C_3H_6N_6O_6$	222.1	1.82	+6.61	2.701
HMX	$C_4H_8N_8O_8$	296.2	1.90	+6.05	2.701

SMOKY - Propellants containing metals (such as aluminum) give exhaust with particulate matter (such as aluminum oxide) which appears in the exhaust stream as smoke. Solid exhaust products such as aluminum oxide are called primary smoke. Smoky propellants formulated to reduce HCl emissions to less than one percent of the exhaust gas mixture are termed clean propellants.

REDUCED SMOKE - Propellants without metals or primary smoke, but containing oxidizers such as ammonium perchlorate, which gives HCl gas as a principal combustion product, are called reduced smoke propellants. If HCl is exhausted in atmospheres of high or moderate humidity, water droplets will coalesce about the HCl molecules, resulting in a visible exhaust trail of what is called secondary smoke. In atmospheres of low humidity, the exhaust plume of reduced smoke propellants is not visible.

MINIMUM SMOKE - Propellants with no metals, and having exhaust free of nucleating species such as HCl, are termed minimum smoke propellants.

MINIMUM SIGNATURE - Propellants whose exhaust characteristics are tailored to give not only minimum smoke properties, but also to have low visible, ultraviolet, or infrared emissions are termed minimum signature propellants. Minimum signature propellants are of interest from the standpoints of launch site and missile detectability and from considerations of through-plume guidance.

Fig. 30.16. Classification of propellants based on exhaust properties.

storage has resulted in recent emphasis on the development of insensitive (low hazard) rocket propellants for tactical applications.

A number of other materials have received attention as potential oxidizers for propellant use, but to date have found little actual use. They include hydroxylammonium nitrate (HAN), hydroxylammonium perchlorate (HAP), hydrazinium nitrate (HN), and hydrazinium perchlorate (HP).

Metallic Fuels. In rocket applications where exhaust smoke is not a major concern, the use of metallic fuels adds considerable impulse to the composition. By far the most common metal in use as a solid propellant fuel is aluminum because of a combination of several desirable properties:

- Low equivalent weight
- High heat of formation of its oxide
- Low reactivity
- Relatively high density
- Low volatility
- Low cost
- Low exhaust product toxicity

An interesting comparison of some properties of metals with respect to their use as solid propellant fuels is available[51].

Although theoretical considerations indicate that boron and beryllium might be preferred to aluminum, practical considerations dictate otherwise. Some of the theoretical advantage of boron is lost because of the volatility of the oxide, and because boron is oxidized to a mixture of oxidation states, not cleanly to the trivalent oxide. The use of beryllium is, in general, not possible because of the high toxicity of the metal and its exhaust products. Aluminum is preferred to magnesium because of its lower equivalent weight and reactivity; aluminum metal powder normally has a thin oxide coating that diminishes its reactivity until combustion temperatures are reached.

Binder. The binder of a composite solid propellant serves the dual function of providing a matrix to hold the oxidizer and metal fuel and of serving as a fuel itself. In early years, polysulfide-based organic polymers and acrylate polymers were employed as binders, but they have been largely supplanted by polymers based on polybutadiene, polypropylene glycol, polyethylene glycol, or polyesters. The two most important types of pre-polymers used in present propellants are those terminated with carboxyl or with hydroxyl functional groups. These functional groups are used to react with curing agents (cross-linkers), as shown in Fig. 30.17, to provide the high molecular weight polymers that function as propellant binders. In general, the lower molecular weight pre-polymer introduced into the propellant formulation is di- or polyfunctional so that the polymer resulting from reaction with a curing agent has a degree of cross-linking sufficient to lend a desirable degree of rigidity to the flexible propellant. The most common pre-polymers in recent propellant use are the hydroxy-terminated polybutadienes; in the previous generation, carboxy-terminated polybutadienes were employed. For example, the binder used in the space shuttle solid booster propellant is based on PBAN, a carboxy-terminated terpolymer of butadiene, acrylic acid, and acrylonitrile. Hydroxy-terminated polyethers also are used, particularly in high performance propellants. Hydroxy-terminated pre-polymers, which are cured with isocyanates to give polyurethanes, offer advantages over carboxylic acid terminated pre-polymers:

- Lower mix viscosity
- Faster, lower temperature cure
- Lower susceptibility to side reactions

Plasticizers. In general, propellant formulations include plasticizers, which are non-reactive diluents used to improve processing and mechanical properties (particularly the low temperature properties) of the propellant. Plasticizers such as high-boiling esters (e.g., dioctyl adipate) or low molecular weight isobutene oligomers frequently are used. In energetic formulations, nitrato ester plasticizers, such as NG, are used not only to improve processing and low temperature properties but also to improve impulse and to serve as oxidizer. In such formulations, high plasticizer/polymer ratios (2–3) frequently are used. The nitrato ester plasticizers are not compatible with the butadiene-based polymers. The binder of the propellant usually is considered to consist of the polymer, the curing agent, and the plasticizer.

Other Propellant Chemicals. In addition to the binder, oxidizer, and fuel, a solid propellant

Fig. 30.17. Curing reactions used in present propellants. Each reactant is di- or polyfunctional, so that high molecular weight polymers are formed as the propellant binder.

may have a variety of other chemicals added (usually in small amounts) for specific purposes. These include:

- Aging stabilizers
- Processing aids
- Bonding agents
- Cross-linking agents
- Burning rate modifiers
- Signature-modifying agents
- Cure catalysts
- Combustion stability enhancers

The final propellant composition is a result of the interaction of a considerable number of chemicals, each of which is important and is selected for one characteristic or for several reasons. The ultimate purpose of the formulation is to give a propellant whose properties are reproducible from batch to batch and from motor to motor, and are adequate for the intended use as shown, for example, in Fig. 30.18.

Propellant Use Criteria

To function properly in its intended use, a propellant must satisfy a large number of criteria, as discussed in the following paragraphs.

Performance. The composition must have adequate specific impuse and volumetric impulse to perform its mission. Volumetric impulse is the product of impulse and density (or density raised to a fractional power).

Mechanical Properties. The important propellant properties include the tensile strength, strain capability (elongation), modulus of elasticity, and strain endurance. The propellant is formulated so that it will be sufficiently

Fig. 30.18. The solid booster rockets for the space shuttle are one of the most widely publicized applications of solid rocket propellant.

flexible to withstand the stresses of acceleration and temperature changes without cracking, yet be sufficiently rigid so as not to slump or deform upon standing or undergoing temperature changes. Bonding agents, which improve the interaction between polymer and filler (oxidizer), frequently are employed to improve mechanical properties.

Bond Properties. The strength of the propellant-to-liner-to-insulation bond must be sufficient to maintain its integrity under the stresses mentioned above.

Ballistic Properties. Important parameters in this regard are the burning rate, the pressure dependence of the burning rate, and the temperature dependence of the burning rate. The burning rate is adjusted using the oxidizer particle size and combustion-modifying additives as variables. With AP oxidizer, finer oxidizer particles give faster burning rates. Finely divided iron oxide is a catalyst frequently employed to accelerate the burning rate of AP propellants. In general, low sensitivity of the burning rate to changes in pressure or temperature is mandatory. Burning rates and ballistic properties are measured in progressively larger motor firings as development of a propellant proceeds.

Combustion Stability. Although burning without acoustic oscillations is partly a motor design concern, the propellant may be modified by the addition of refractory particles to dampen such vibrations. If uncorrected, pressure oscillations from combustion instability may be large enough to destroy a motor during firing.

Aging and Service Life. Propellants must be storable for at least as long as the intended service life of the missile system without undergoing degradation or change of ballistic or mechanical properties. Usually the service life is estimated by extrapolation from properties measured for samples aged at elevated temperatures. This accelerated aging is presumed to speed processes that would take place at lower (storage) temperatures. Stabilizers selected to enhance the aging capabitity are usual components of propellant compositions; the stabilizer choice depends on the polymer, plasticizer, and oxidizer types.

Processibility and Castability. In order to be introduced into a rocket motor, a castable propellant composition must be blended until all components are evenly dispersed. The resulting composition must be sufficiently fluid that it may be cast into the motor without creating voids or bubbles in the propellant and yet be sufficiently viscous that dense particles (oxidizer, aluminum) do not settle, or less dense materials (polymer) rise to the surface. Certain chemicals can significantly improve the processibility of propellants when added in small amounts.

Potlife and Cure. In addition to achieving a castable viscosity, it is necessary that the liquid (uncured) propellant remain fluid for a sufficient amount of time to be transported and cast, with allowance for possible delays, before it solidifies appreciably. Following cast, the composition must be cured to a solid, preferably at temperatures close to the intended storage temperature. The balancing of potlife and cure rate frequently requires careful adjustment of cure catalyst levels, mix temperature, and cure temperature.

Hazard Properties. It must be verified that the propellant is sufficiently insensitive to shock, electrostatic discharge, friction, thermal decomposition, or self-heating (in larger quantities) that it does not represent an unwarranted hazard in its intended use. Rocket propellants are energetic compositions and must be formulated so that chance stimuli will not initiate violent reaction.

Ignitability. Conditions for ignition in the desired application must be defined, and the propellant formulated so that it may be reliably ignited under these conditions.

All of the above factors must be carefully studied and optimized before a solid propellant can be considered adequate for its intended use.

Composite Propellant Manufacture

In a typical batch processing sequence for a polyurethane composite propellant, a submix is prepared first. The submix contains the pre-polymer, plasticizer(s), stabilizers, and usually a bonding agent; it is slurried with aluminum powder to give a premix. The premix is added to a moderate-shear slurry mixer, and ammonium perchlorate oxidizer is added in several portions with intervening mixing. Following addition of the oxidizer, the composition is mixed under vacuum for a defined period. Vacuum is released, and the curing agent and cure catalyst(s) are added. After a final vacuum mix, the propellant is ready for casting. Frequently, the propellant is cast into an evacuated motor; the use of vacuum increases the casting rate and lessens the possibility of air entrapment. The cast motor then is placed in a cure oven (typically held at a temperature in the range of 110–160 °F) until the propellant has cured. At many points in the process, samples are taken and analyzed to ensure quality.

Although most solid propellants are manufactured in a batch process, a continuous mixing process has been used successfully in the production of the first stage A-3 Polaris propellant. The use of a continuous mixing process, in which propellant chemicals are metered into a helical kneader, offers considerable benefit in safety and cost for large-volume propellant production.

LIQUID PROPELLANTS

To date the liquid propellant systems used in chemical propulsion range from a small trajectory control thruster with only 0.2 lbf (0.89 N) thrust to large booster rocket engines

with over 1.0 million lbf (4.44 MN) thrust. Bipropellant propulsion systems are the most extensively used type today for applications in main combustion chambers and gas generators. The monopropellant propulsion system is widely used in low temperature gas generators and auxiliary rockets for trajectory or orbital adjustment.

A major difference between liquid propellants and solid propellants used in chemical propulsion systems is the ease of use or controllability. The solid propellants are cast as a solid propellant grain. The burning rate is dependent on the propellant formulation and the configuration design of the solid propellant grain in addition to the chamber pressure and grain temperature. The combustion process is continuous, and a quench and reignition combustion process is difficult. For the liquid bipropellant system, the liquid fuel and the liquid oxidizer are stored in separate tanks and fed separately to the combustion chamber. The propellants are fed either by means of pumps or by pressurization with an inert gas. A controller generally is used to control the flow rate of the liquid propellants in the system. Ignition and reignition combustion generally is employed to fit the mission requirements and objectives.

Fig. 30.19. Liquid bipropellant rocket engine.

An injector is used to atomize and mix the liquid propellants in appropriate proportions. The propellants enter the thrust chamber through the injector and burn inside the thrust chamber. A typical liquid bipropellant rocket engine is shown in Fig. 30.19.

Liquid propellants can be categorized by their type of storage (cryogenic propellants vs. storable propellants) or by their function in the chemical propulsion system (oxidizers, fuels, or monopropellants). The functional categorization is used here for discussion of the liquid propellants. In addition, liquid propellants, either oxidizer or fuel, can be gelled as a neat material or as a heterogeneous gel mixture containing suspended solid material such as aluminum powder.

Physical Properties

General physical properties, including freezing point, normal boiling point, critical temperature, critical pressure, specific gravity, heat of formation, and heat of vaporization, of several common liquid propellants are listed in Table 30.8. Of the temperature-dependent physical properties, such as heat capacity, thermal conductivity, viscosity, and specific gravity, only the specific gravity is included in Table 30.8. Generally, cryogenic propellants are listed at their normal boiling point, whereas the storable propellants are evaluated at 68 °F (293 °K). Detailed information on the physical properties of the liquid propellants can be found in reference 52. To obtain a wide operating range and a large payload capacity, the desired physical properties are:

- Low freezing point
- Low temperature variability
- Low vapor pressure
- High specific gravity
- High heats of formation and vaporization

TABLE 30.8 Properties of Common Liquid Propellants

Propellant	*Molecular Weight*	*Freezing Point (°F)*	*Boiling Point (°F)*	*Critical Temp. (°F)*	*Critical Press. (psia)*	*Specific Gravity*	*Heat of Formation (cal/mole @298.16°K)*	*Heat of Vapor. (Btu/lb :NBP)*
LO_2	32.0	−362	−297	−182.0	730.6	1.14[a]	−2896	91.62
F_2	38.0	−365	−307	−201	808.5	1.50[a]	−3056	71.5
N_2O_4	92.011	11.75	70.4	316.8	1441.3	1.45[b]	−4.7	178.2
ClF_5	130.445	−153.4	7.3	289.4	771	1.795[b]	−60,500	76.04
H_2O_2	34.016	31.2	302.4	855	3146	1.38[b]	−44,750	76,04
H_2	2.016	−434.8	−423.3	−399.9	188	0.071[a]	−1895	195.3
N_2H_4	32.045	34.75	237.6	716	2131	1.008[b]	12,054	583
MMH	46.072	−62.3	189.8	594	1195	0.879[b]	13,106	377
UDMH	60.099	−70.94	144.18	482	867	0.785[b]	12,339	250.6
A-50 (50% N_2H_4—50% UDMH)	41.802	22.0	158	633	1731	0.905[b]	12,310	346.5
RP-1(H/C = 2.0)	172.0	−55	422	758	340	0.807[b]	−6222	125
H_2O	18.016	32	212	705.4	3206.2	1.0[b]	−68,317	970.3

[a]Evaluated at NBP.
[b]Evaluated at 68°F (293.4°K).

Because liquid propellants also are used to cool the thrust chamber assembly, good heat transfer properties, such as high heat of vaporization, high thermal conductivity, high specific heat, and high boiling point, are desirable.

Liquid Oxidizers

Liquid oxidizers can be categorized as either storable or cryogenic. Many different types of liquid oxidizers have been used in chemical propulsion systems. In general, cryogenics such as liquid oxygen, fluorine, or fluorinated compounds give a high specific impulse. Several storable oxidizers such as nitrogen tetroxide (N_2O_4), inhibited red fuming nitric acid (IRFNA), or chlorine pentafluoride (ClF_5) have been used for their advantages in storage. A brief description of the commonly used oxidizers is given below.

Cryogenics. Liquid oxygen, the most important and extensively used liquid oxidizer, is used primarily with liquid hydrogen to give a very high specific impulse (usually near or over 400 lbf-sec/lbm). Major applications of the liquid oxygen/liquid hydrogen bipropellant system include the space shuttle main engine and the Saturn second stage engine (J-2). Liquid oxygen also is widely used with hydrocarbon fuels in the booster engine of heavy lift launch vehicles such as the Russian Energia. Although liquid oxygen can be used with storable fuels, such as hydrazine or monomethylhydrazine (MMH), this bipropellant combination engine is still in the development stage, mainly because of the difficulty of chamber cooling and combustion stability.

Both the freezing point (−362 °F, 54.5 °K) and the boiling point (−297 °F, 90 °K) of liquid oxygen are low, permitting a wide range of operation. Liquid oxygen is highly reactive with most organic materials in a rapidly pressurized confinement. Except for the relatively high evaporation rate, the handling and storage problems for liquid oxygen are minimal. Although liquid oxygen is not considered corrosive and toxic, the surfaces that will be in contact with the liquid oxygen must be kept extremely free of any contamination because of its reactivity. The low boiling point also can cause problems due to low temperature embrittlement. In order to minimize oxidation problems, metals such as copper should be used.

Storage tanks and transfer lines of liquid oxygen systems must be well insulated to prevent the condensation of moisture or air with subsequent ice formation on the outside. Vacuum jackets, formed plastics, and alternate layers of aluminum foil and glass-fiber mats have been used successfully.

Liquid fluorine and fluorine compounds (F_2, OF_2, or NF_3) are also cryogenic oxidizers. Although fluorine offers specific impulse and density advantages, extreme toxicity and corrosiveness have prevented the practical application of fluorine and fluorinated compounds in chemical rocket design. In addition, with fluorine the handling problems are significant because fluorine has the highest electronegativity of any of the elements and a low boiling point (85.0 °K).

Storable. Nitrogen tetroxide and IRFNA are the most important and most extensively used storable liquid oxidizers. A high density yellow-brown liquid, nitrogen tetroxide is very stable at room temperature, but at 303 °F (423 °K) it begins to dissociate into nitrogen dioxide. Upon cooling, nitrogen tetroxide is re-formed. It is used as the oxidizer for the Titan first and second stage rocket engines, Delta second stage rocket engine, and for the orbital maneuver engines of the space shuttle.

The freezing temperature of nitrogen tetroxide (11.75 °F, 262 °K) is relatively high; so care must be taken to avoid freezing whenever it is used. Nitrogen tetroxide is not corrosive if pure, that is, if the moisture content is very low. Carbon steels, aluminum, stainless steel, nickel, and Inconel can be used with it. However, when the moisture content is high, the nitrogen tetroxide becomes corrosive, and 300-series stainless steel should be used.

IRFNA consists of concentrated nitric acid (HNO_3) that contains dissolved nitrogen dioxide and a small amount of fluoride ion.

The addition of HF provides the fluoride ion, which reacts with the metal containers, forming a metal fluoride coating and reducing the corrosiveness significantly. As with nitrogen tetroxide, IRFNA is mainly used with hydrazine-type fuels in bipropellant systems. Because nitric acid ignites spontaneously with aniline and amines, care must be taken in handling and storing the IRFNA. However, this is an advantage in systems where extinguishment and reignition are desired.

Mixed oxides of nitrogen (MON) and high density acid (HDA) are variants of storable oxidizers. MON consists of N_2O_4 and NO, and HDA is a mixture of nitric acid and N_2O_4. Both HDA and MON behave similarly to IRFNA.

Other liquid oxidizers such as hydrogen peroxide (H_2O_2), chlorine trifluoride (ClF_3), and chlorine pentafluoride (ClF_5) are used as rocket propellants for special applications. Hydrogen peroxide seldom is used because of its poor storage stability. Chlorine trifluoride (ClF_3) and ClF_5 can be highly corrosive to metals. Other compounds such as ClO_3F, ClF_3O, or BrF_5 also can be considered as alternatives.

Liquid Fuels

Cryogenics. Liquid hydrogen, the most important and widely used liquid fuel, is used mainly with liquid oxygen to give high performance, as mentioned above. Liquid hydrogen has excellent heat transfer characteristics: high heat of vaporization, high specific heat, and high thermal conductivity. It is a very good choice to cool the thrust chamber assembly when it is used in a regeneratively cooled system. However, the low boiling point (−434.8 °F, 14 °K) and low density of liquid hydrogen make it difficult to handle and store. The low fuel density means that a very large and bulky fuel tank is required; this is considered a disadvantage when a high ratio of payload to vehicle dry weight is desired. Liquid hydrogen has been found to react with metals, causing embrittlement of such metals as nickel. Copper is the best and most widely used material for liquid hydrogen in rocket engine applications.

As with liquid oxygen, storage tanks and transfer lines of liquid hydrogen systems must be well insulated, to prevent the evaporation of hydrogen or condensation of moisture or air with subsequent ice formation on the outside. As mentioned before, vacuum jackets, formed plastics, or glass-fiber mats mixed with aluminum foil can minimize the problems.

Molecular hydrogen exists in two forms: *ortho*-hydrogen (nuclei of the two atoms spinning in the same direction) and *para*-hydrogen (nuclei of the two atoms spinning in opposite directions). These two forms are in equilibrium with each other, and at room temperature the equilibrium mixture contains 75 percent of the *ortho* form and 25 percent of the *para* form. Upon cooling to the normal boiling point, the equilibrium is shifted. At this temperature, the *ortho* form will convert slowly to the *para* form. The equilibrium concentration of *para*-hydrogen at this temperature, the normal boiling point, is 99.8 percent. Therefore, *para*-hydrogen generally is used as liquid fuel for rocket engine applications. Modern liquefiers can produce liquid hydrogen that is more than 99 percent *para*-hydrogen.[53]

Liquid methane (CH_4), another kind of cryogenic fuel used in testing and experiment, generally is used with liquid oxygen in a bipropellant system. It has the advantage of a considerably higher density than that of hydrogen (the specific gravity for methane is 0.4507 at the normal boiling point). To date no operational rocket engine utilizes liquid methane as the fuel; all liquid oxygen/liquid methane engines are still in the development stage.

Storables. Hydrazine-type fuels are the most important storable liquid fuels. They include hydrazine (N_2H_4), monomethyl hydrazine (MMH), unsymmetrical dimethylhydrazine (UDMH), Aerozine-50 (50% N_2H_4 and 50% UDMH), and various blends of these fuels with other components. All hydrazine-type fuels are toxic to some degree.

A colorless liquid, hydrazine is stable to shock, heat, and cold. The freezing point of hydrazine (34.75 °F, 274.9 °K) is the highest of commonly used hydrazine-type fuels. Because it starts to decompose at 320 °F (433 °K) with no catalysts present, it is undesirable for use as the coolant for regenerative cooling of the thrust chamber. Different blends of hydrazine and MMH have been tested to improve heat transfer properties. Hydrazine is generally compatible with stainless steel, nickel, or aluminum. (See Chapter 28 for more information on hydrazine.)

UDMH is neither shock- nor heat-sensitive, and it is a stable liquid even at high temperature. Furthermore, it is compatible with nickel, Monel, and stainless steel.

UDMH is often used as rocket propellant mixed with hydrazine in various proportions. A 50/50 mixture, Aerozine-50, is used for the current Titan IV engines.

MMH generally is used with nitrogen tetroxide (N_2O_4) in small spacecraft rocket engines such as orbital maneuver or altitude control engines. Compared to hydrazine, MMH has better shock resistance and better heat transfer properties as a coolant. However, the specific impulse for MMH/N_2O_4 engines is slightly lower than that for hydrazine/N_2O_4 bipropellant engines. Like hydrazine, MMH is compatible with stainless steel, nickel, aluminum, Teflon, and Kel-F.

In general, the hydrazine-type fuels do not have very good heat transfer properties. Therefore, in the latest development of a high pressure bipropellant system using N_2O_4 and hydrazine-type fuels, the oxidizer N_2O_4 has been used as a regenerative coolant instead of the fuel itself.

Monopropellants

Simplicity and low cost are the major reasons why monopropellant rocket engines are considered for development. The specific impulse for monopropellant engines generally is much lower than that for bipropellant engines (generally, in the range of 200 lbf-sec/lbm for monopropellant vs. 280–400 lbf-sec/lbm for bipropellant). Hydrazine is the most important and widely used monopropellant in small trajectory correction or altitude control rockets. In an effort to lower the freezing point, many hydrazine blends have been studied.

Hydrogen peroxide has been used as a monopropellant, especially in various solutions with water. It was used as a rocket propellant in the X-15 research aircraft. However, because of storage stability problems and low performance, hydrogen peroxide now is no longer used.

Gelled Propellants

As with metallic fuels in solid composite propellants, early interest in gelled liquid propellants focused on gelation as a means of incorporating high energy solids into propellants to achieve high specific impulse and density. Storable hydrazine-type fuels such as hydrazine and MMH have received most of the attention in gelation development. During the development of Titan IIA, a stable suspension of aluminum powder in gelled hydrazine (called Alumazine) was developed and tested extensively. Later the needs for such a vehicle disappeared, leaving the gelation technology to languish for lack of other applications.

In the early 1980s, an increasing focus on improving the safety and handling characteristics of storable liquid propellants revived the interest in developing gelled liquid propellants. Although gelation technology for both cryogenic and storable liquid propellants has been developed, most of the activities have concentrated on storable propellants, especially in liquid fuels. Comprehensive reviews covering all gelled fuels development activities can be found in references 54 and 55. A brief description of the gelled fuels and oxidizers is given below.

Gelled Fuels. Gelled fuels generally have consisted of metallic powders such as aluminum or boron suspended in gelled MMH or an MMH blended fuel. These gels are applied to tactical missile applications. The gelled

fuels without metal additives have drawn added interest as propellants with minimum exhaust signatures and relatively high specific impulse.

Gelled fuels have essentially the same toxicities as their ungelled counterparts because gelation does not significantly change their compositions or equilibrium vapor pressure. However, the rates of vaporization are decreased significantly, reducing the toxic exposure hazards. Gelled fuels, such as gelled MMH, are hypergolic but can burn only upon direct contact of the fuel gel with oxidizer; the fire hazards are greatly reduced with respect to both intensity and extent. Compatibilities of the gelled fuels generally are the same as those of the ungelled fuels.

Gelled Oxidizers. Development work on gelled oxidizers was started during the 1960s, but virtually no effort was expended on them during the 1970s. Development work was revived in the 1980s as interest in gels was awakened because of their improved safety and handling characteristics. Gelled oxidizers include nitrogen tetroxide, mixed oxides of nitrogen (MONs), red fuming nitric acids (RFNA), and IRFNA. The gelling agents used for the oxidizers include powdered lithium nitrate ($LiNO_3$), lithium fluoride (LiF), plastisol nitrocellulose (PNC), and colloidal silica. The early work on these gels can be found in references 56 and 57. For tactical applications with minimum exhaust signatures, gelled IRFNA and ClF_5 are the primary gel oxidizers.

Like the gelled fuels, the toxicity of gelled oxidizers is not reduced, but the handling and safety hazards of storage are greatly improved. Gelled IRFNA is considered to be compatible with aluminum. Recent efforts have shown that it can be stored for at least as long as neat IFRNA.

REFERENCES

1. Urbanski, T., *Chemistry and Technology of Explosives*, Vol. I, Pergamon Press, Elmsford, New York, 1964.
2. Cook, M. A., *The Science of High Explosives*, Van Nostrand Reinhold, New York, 1958.
3. Compiled from annual issues of *Mineral Industry Surveys*, Bureau of Mines, U.S. Department of the Interior, Washington, DC ("Apparent Consumption of Industrial Explosives and Blasting Agents").
4. Van Gelder, A. P., and Schlatter, H., *History of the Explosives Industry in America*, Arno Press, New York, 1972 (reprint of 1927 edition).
5. Dannenberg, J., *Proceedings of the Sixth Conference on Explosives and Blasting Technique*, p. 469, Society of Explosives Engineers, 1980.
6. Ornellas, *J. Phys. Chem.*, **72**, 2390 (1968).
7. Kaye, S. M., *Encyclopedia of Explosives and Related Items*, PATR 2700, Vol. 8, pp. P99–100, U.S. Army Armament Research and Development Command, 1978.
8. *Engineering Design Handbook: Explosives Series*, Army Material Command Pamphlet 706-177, AD 764,340, p. 12, distributed by NTIS, Jan. 1971.
9. Cook, M. A., *The Science of High Explosives*, p. 14, American Chemical Society Monograph Series No. 139, 1958.
10. *Ibid.*, p. 315.
11. Sheahan, R. M., and Minchinton, A., *The Australian IMM Explosives in Mining Workshop*, p. 17, Melbourne, Nov. 1988.
12. Fedoroff, B. T., and Sheffield, O. E., *Encyclopedia of Explosives and Related Items*, PATR 2700, Vol. 2, p. B165, Picatinny Arsenal, Dover, NJ, 1962.
13. Urbanski, T., *Chemistry and Technology of Explosives*, p. 322, Vol. III, Pergamon Press, New York, 1967.
14. "Properties of Explosives of Military Interest," *Engineering Design Handbook: Explosives Series*, AD 764,340, p. 38, NTIS, U.S. Department of Commerce, 1971.
15. Butler, A. R., *ChemTech*, p. 202 (Apr. 1990).
16. Marshall, A., *Explosives*, Vol. I, *History and Manufacture*, P. Blakiston's Son & Co., Philadelphia, PA, 1917.
17. Gregory, C. E., *Explosives for North American Engineers*, Trans Tech Publications, Cleveland, OH, 1973.
18. "Black Powder," *Explosives and Pyrotechnics*, **16** (7) (July 1983).
19. *Explosives and Rock Blasting*, Atlas Powder Co., Dallas, TX, 1987.
20. *Blasters Handbook*, 16th ed., pp. 14 and 109, Du Pont, 1980.
21. Johansson, C. H., and Persson, P. A., *Detonics of High Explosives*, Academic Press, New York, 1970.

22. Kintz, G. M., et al., Report of Investigations 4245, U.S. Dept. of Interior, Bureau of Mines (Feb. 1948).
23. *Chemical Marketing Reporter*, p. 6 (Sept. 19, 1983).
24. Cook, M. A., *The Science of Industrial Explosives*, p. 2. IRECO Chemicals, Salt Lake City, UT, 1974.
25. Egly, R. S., and Neckar, A. E., U.S. Patent 3,161,551, Dec. 15, 1964; Bluhm, H. F., U.S. Patent 3,447,978, June 3, 1969.
26. Bower, J. K., et al., *I & EC Product Research & Development*, p. 326 (Sept. 1980).
27. Mohan, V. K., and Field, J. E., *Combustion and Flame*, **56**, 269 (1984).
28. Spear, R. J., and Wilson, W. S., *J. Energetic Materials*, **2**, 61 (1984).
29. Coburn, M. D., et al., *Ind. Eng. Chem. Prod. Res. Dev.*, **25**, 68 (1986).
30. Dobratz, B. M., LA-9732-H, VC-45, Los Alamos National Laboratory, Los Alamos, NM (May 1983).
31. Cranney, D. H., and Hales, R. H., *Proceedings of the Fourteenth Symposium on Explosives and Pyrotechnics*, Franklin Research Center, Feb. 1990.
32. Seltzer, R. J., *Chemical and Engineering News*, p. 7 (Aug. 8, 1988).
33. Code of Federal Regulations, Title 49 Transportation, Part 173.114a, p. 272, Oct. 1, 1979 Revision.
34. *Federal Register*, **55** (9), 1306 (Jan. 12, 1990).
35. Tarver, C. M., et al., "Structure/Property Correlations in Primary Explosives," Final Report, 76-2, Stanford Research Institute, Menlo Park, CA (Feb. 4, 1977).
36. Brunswig, H., *Explosivstoffe*, p. 17, Barth, Braunschweig, 1909, and Plets, V., *Zh. Obshch. Khim*, **5**, 173 (1953), as cited in reference 1, p. 1.
37. Shipp, K. G., *J. Organic Chemistry*, **29**, 2620 (1964); U.S. Patent 3,505,413, Apr. 7, 1970.
38. Kayser, E. G., *J. Energetic Materials*, **1**, 325 (1983).
39. Gallo, A. E., and Tench, N., *J. Hazardous Materials*, **9**, 5 (1984).
40. Delistraty, J., and Brandt, H., *Propellants, Explosives, Pyrotechnics*, **7**, 113 (1982).
41. Rizzo, H. F., et al., *Propellants and Explosives*, **6**, 27 (1981); Kolb, J. R., and Rizzo, H. F., *Propellants and Explosives*, **4**, 10 (1979).
42. Ott, D. G., and Benziger, T. M., *J. Energetic Materials*, **5**, 343 (1987).
43. Urbanski, T., *Chemistry and Technology of Explosives*, Vol. II, p. 181, Pergamon Press, New York, 1965.
44. *Military Explosives*, Dept. of the Army Technical Manual TM 9-1300-214, p. 7–32, Washington, DC, 1967.
45. Watson, S. C., and Winston, S. E., *Proceedings of the Eleventh Conference on Explosives and Blasting Technique*, p. 320, Society of Explosive Engineers, 1985.
46. Sudweeks, W. B., *I & EC Product Research & Development*, **24**, 432 (1985).
47. *Work Injury & Illness Rates*, pp. 6 and 8, National Safety Council, Chicago, IL, 1988.
48. Sutton, G. P., *Rocket Propulsion Elements*, 5th ed., Wiley-Interscience, New York, 1986.
49. Gordon, S., and McBride, B. J., *Computer Program for Calculation of Complex Chemical Equilibrium Compositions. Rocket Performance, Incident and Reflected Shocks, and Chapman–Jouget Detonations*, NASA SP-273, Interim Revision, NASA Lewis Research Center, (Mar. 1986).
50. Oberth, A. E., *Principles of Solid Propellant Development*, CPIA Publication 469, Chemical Propulsion Information Agency, Baltimore, MD, 1987.
51. Dekker, A. O., "Solid Propellants," *J. Chem. Ed.*, **37**, 597.
52. Cadwallader, A., *Liquid Propellant Manual*, Chemical Propulsion Information Agency, 1969.
53. Von Doehern, J., *Propellant Handbook*, AFRPL-TR-66-4, Air Force Rocket Propulsion Laboratory, 1966.
54. Vander Wall, E. M., Anderson, R. E., and Cabeal, J. A., "Preparation and Characterization of Gelled Hydrazine Based Fuels Containing Aluminum," Final Report, AFRPL-TR-80-40, Air Force Rocket Propellant Laboratory (July 1980).
55. Haun, D. V., Ryder, D. D., and Graham, P. H., "Gelled Bipropellants for Advanced Missile Propulsion Systems, I. Fuel Component Initial Characterization," CPIA Publication 425, Vol. VI, pp. 215–224, JANNAF Propulsion Meeting, Apr. 1985.
56. Vander Wall, E. M., Anderson, R. E., Beegle, R. L., Jr., and Cabeal, J. A., "Gelation of N_2O_4 and MON-25," Final Report, AFRPL-TR-82-49, Air Force Rocket Propulsion Laboratory (Aug. 1982).
57. Allan, B. D., "A Gelled Oxidizer for Tactical Missiles," CPIA Publication 370, Vol. V, pp. 11–20, JANNAF Propulsion Meeting, Feb. 1983.

31

The Nuclear Industry

William M. Eister* and Warren K. Eister**

INTRODUCTION

The objective of the nuclear industry is to produce energy in the forms of heat and radiation. For heat applications, the nuclear fuel has a very high specific energy content that has two principal uses, for military explosives and for electricity generation. For radiation applications the emissions from radioactive decay of unstable nuclides are employed in research, medicine, and industry for diagnostic purposes and for chemical reaction initiation.

This nuclear technology is based on both the nuclear and the chemical properties of the atom. At the beginning of the twentieth century fewer than 90 chemical elements were known, and there was only a dawning awareness of isotopes. Today, largely because of the nuclear industry, approximately 2500 isotopes (or nuclides, depending on the properties of interest) and 470 isomers have been identified. This includes 180 nuclides and 40 isomers identified between 1983 and 1988. Brief definitions of several chemical and nuclear terms are given in Table 31.1.

Like other technology, nuclear technology involves a combination of science and art. However, it is unique because of the development of the atomic bomb that contributed to the ending of World War II. Many people view nuclear technology as a combination of magic and illusion that is difficult to separate from reality, both inside and outside the industry.

TABLE 31.1 Definitions of Atoms, Chemical Elements, Isotopes, Nuclides, and Isomers

Atoms—elementary particles of matter.
Chemical elements—atoms with unique properties related to their orbital electrons.
Isotopes—atoms of the same chemical element with different masses related to a different number of neutrons in the nucleus.
Nuclides—atoms with unique properties related to the neutrons and protrons in the nucleus of the atom.
Isomers—nuclides with different energy states.

*Oconee Nuclear Power Station, Duke Power Company.
**Formerly with the Oak Ridge National Laboratory and the U.S. Department of Energy.

Fig. 31.1. Uranium metal fuel from the first atomic pile. This reactor contained 40 tons of uranium oxide along with 6.2 tons of uranium metal. (*ORNL News, 1-01-076.*)

Fig. 31.2. The nuclear industry.

This chapter first considers the following subjects:

- The status and outlook of the nuclear industry.
- Nuclear safety.
- The role of nuclear energy in meeting the world's energy needs.
- The nuclear processes on which the nuclear industry is based (Fig. 31.2a).
- The fuel cycle for nuclear electric power generation, which involves activities typical of the chemical process technologies (Fig. 31.2b).

This is followed by discussions of radioactive waste management and transportation of nuclear materials, and consideration of various applications, such as nuclear power reactors used for the generation of electric power, the use of radioisotopes, and other military and civilian uses of nuclear materials (Fig. 31.2c).

A reference section is provided for those who wish to have more detailed information.

TABLE 31.2 Nuclear Industry Products

	Approximate Number
Civilian:	
Electric generating plants	450
Ships (ice breakers and transports)	10
Radioisotope applications (tracers, radiation sources, thermal sources)	Millions
Military:	
Weapons	30,000
Submarines	200
Ships (aircraft carriers, cruisers, transports)	30

STATUS AND OUTLOOK

The nuclear industry provides products that play vital societal roles:

- As nuclear explosives and in nuclear-powered submarines and ships, which have contributed to maintaining world peace since World War II.
- As a primary energy source for the production of electricity to meet the world's energy needs and without polluting the atmosphere.[1]
- As stable and radioactive nuclides, which contribute significantly to research, medicine, and industry (see Table 31.2 and Fig. 31.2.)

Nuclear technology continues to be developed for military applications, and nuclear explosives, submarines, and ships continue to be produced. Brief attention will be given herein to explosives, primarily in terms of their historical role in the excavation of harbors and canals and in the enhancement of gas and petroleum deposits.

Within the past half-century, since the discovery of nuclear fission, this technology has been developed to supply 17 percent of the world's electricity from 450 nuclear reactors located in 26 countries.[2] This is an international program regulated by the United Nations' International Atomic Energy Agency, which includes 113 member countries. In 1990 there were 111 nuclear units in the United States, including 38 boiling water reactors, which together generated 570 billion kWh of electricity, some 20 percent of the country's total electricity generation. This represented a 127 percent increase in output since 1980. (See Fig. 31.3.)

The safety record of the nuclear industry has continued to be generally very good, despite the 1987 accident with a nuclear reactor at Chernobyl, in the Ukraine, which killed a number of workers, caused the abandonment of adjacent communities and farmland, and resulted in radioactive fallout in neighboring countries. It should be noted that these were the first nuclear fatalities since the beginning of the nuclear power industry.[7] It is generally agreed that this unfortunate occurrence was caused by serious design flaws, and that power reactors in use elsewhere are not subject to similar occurrences. Indeed, a somewhat similar yet also very different accident occurred in the United States in 1979 (Three Mile

Fig. 31.3. The Oconee Nuclear Station, Duke Power Company. The total electrical capacity of its three pressurized water reactors is 2580 megawatts, and its two pumped storage units add 1610 megawatts to its peak capacity. (*Courtesy Duke Power Company.*)

Island); but because of the reactor's engineered safety features (required in the United States and recommended by the International Atomic Energy Agency), there were no injuries and no significant radiation exposure, either to workers or the public.[6]

Worldwide, aggressive public concern about nuclear safety has been delaying the development of waste disposal facilities for high-level wastes. Current low-level waste facilities in the United States are nearly full, and there has been great difficulty in siting new facilities. It now appears that for many years the interim solution for the storage of spent fuel elements will be above-ground storage facilities with appropriate safeguards and security. On the energy side, increased attention to conservation and the ready availability of petroleum from the Middle East have combined to end the urgent need for continued nuclear power development. However, the continuing instability of the oil supply was demonstrated by Iraq's invasion of Kuwait in mid-1990, and the need for a U.S. energy policy that places less reliance on oil has again become apparent.

Light water reactors (LWRs) have been successfully used for electric power generation throughout the world, and heavy water reactors (HWRs) likewise have established good operating records. In recent years significant advances in nuclear reactor technology have been realized. A new generation of light and heavy water reactors has been designed and licensed to further reduce the already very low probability of loss-of-coolant accidents such as occurred at Three Mile Island and Chernobyl. In addition, the enrichment level of new fuel for LWRs has been increased, and the fuel life has thereby been extended from 33,000 to 45,000 megawatt days/metric ton, and the reactor cycle between

shutdowns has been increased from 12 to 18 months.

All the components of the nuclear-fission power system are fully operational except for ultimate waste disposal. However, spent fuel is not reprocessed in the United States because there is an adequate supply of natural uranium available from other countries at a lower cost than that of the recovered fissionable material from spent fuel. Also, the United States unilaterally declared a moratorium on reprocessing to reduce the spread of nuclear weapons.

The nuclear industry makes available about 3000 nuclides, including both the stable and the radioactive nuclides. Approximately 50 radioactive nuclides, along with some stable nuclides that have been isotopically enriched, are essential in research, medical, and industrial applications. Many of these are now produced commercially, but several still are dependent on government facilities. Some, for economic reasons, come from other countries. Radiation processing for sterilization of disposable medical supplies is an important operation using cobalt-60 from Canada. Electron accelerators have replaced radioactive nuclides as radiation sources for polymerizing plastic coatings on wire and paper. Perhaps the greatest disappointment experienced by the industry has been the public opposition to radiation of food to extend its shelf life. The principal current application of this type is by the Russians, who are using electron accelerators for the deinfestation of wheat.

In the continuing research and development of nuclear energy, the liquid metal reactor (LMR) and fusion are the major activities. It now appears that both of these alternatives will significantly increase the cost of electrical energy in comparison with the present LWR/HWR technology based on low-cost uranium ore.

Plant-scale liquid metal reactors are in operation in France and Russia to establish the engineering technology and to evaluate their role in increasing the efficiency of uranium utilization. France, without a significant primary energy supply, wants to minimize imports of uranium. Although the United States imports about 50 percent of its uranium, it continues to depend on its 100-year supply of coal for energy independence. In the United States, engineering test LMRs have operated for many years and have made significant contributions to LMR technology. Whereas they were first studied to reduce natural uranium requirements to approximately 20 metric tons/gigawatt year electrical (GWye), development is now centered on (1) reducing plutonium in the waste, (2) the liquid metal benefit to safety, and (3) electrometallurgical reprocessing at each reactor site, thereby reducing proliferation concern.

It is difficult to assess the progress of fusion research, probably the world's most important international research program. Russia, the European Community, Japan, and the United States are the principal participants. It is now clear that fusion will be a high-cost source of energy and not economically justified within the foreseeable future. However, it is essential for the major world nations to cooperate responsibly in its development. The basic research has yet to achieve controlled ignition of the fusion reaction along with energy parity (energy input less than output).

At this time (1992), the outlook for the nuclear industry is difficult to project. France and Japan, countries without domestic energy resources, are moving ahead to establish efficient nuclear systems for electric generation; and although they have a concerned public, their fears regarding waste disposal and proliferation are not overriding their need for electricity. In the United States, even with abundant coal, it will probably be essential to follow their lead. Otherwise, the United States, once a nuclear pioneer, will become deficient in a significant technology.

NUCLEAR SAFETY

Safety first has always been and continues to be the basic policy of the nuclear industry. This includes activities to discourage the proliferation of nuclear weapons and to prevent sabotage of nuclear facilities. This

poligy has been successful; the chance of death from a nuclear accident is over a million times less than death from ordinary human activities, and over a thousand times less than death from natural events (see Table 31.3). Safety remains the most challenging responsibility of the nuclear industry.

It is the public's strong perception that all nuclear activities are more dangerous than other accepted risks. Much of this public concern results from the atomic bombs and the government's secret program that produced the bombs. However, even with the early large-scale nuclear operations, there were few nuclear fatalities or life-shortening injuries. After 50 years, these facilities are being retired. Meanwhile, considerable operating and safety experience has been gained that very significantly reduce public concern in the future.

The nuclear safety program originally was based on both established laws regulating all industrial safety and early knowledge of radiation health effects related to radium and X-ray exposure. Since then the safety regulations have been greatly enhanced, and they are present in the Code of Federal Regulations, Title 10, Chapter 1, titled Nuclear Regulatory Commission.[5] These regulations strictly control the management, engineering design, and operations of all nuclear activities. In addition the International Atomic Energy Agency, an agency of the United Nations located in Vienna, Austria, has established regulations as a reference for all national programs. Each country also has its own regulatory agency. There are two areas of concern in nuclear safety: (1) radiation exposure of the public and of workers resulting from normal and accident conditions; (2) danger to world peace and order from nuclear proliferation and sabotage.

TABLE 31.3 Average Risk of Fatality by Various Causes in the United States (1975)

Accident Type	*Annual Total Number*	*Individual Annual Risk*
Automobile accidents	56,000	1 in 4,000
Falls	18,000	1 in 10,000
Fires	7,500	1 in 25,000
Drowning	6,200	1 in 30,000
Gun accidents	2,300	1 in 100,000
Air travel	1,800	1 in 100,000
Electricity	1,200	1 in 160,000
Lightning	160	1 in 2,000,000
Hurricanes	94	1 in 2,500,000
Tornadoes	91	1 in 2,500,000
Nuclear reactors	—	1 in 5,000,000,000*
All accidents	112,000	1 in 1,600

*Calculation based on 100 LWRs operating under U.S. Nuclear Regulatory Commission supervision.

Radiation, like air, has a ubiquitous presence in the human environment. Humans are exposed to approximately 100 to 200 millirem/year depending on their location in the world (120 millirem/year at Harrisburg vs. 190 at Denver). The principal natural radiation sources are cosmic rays from outer space, radon from geologic sources, and potassium-40 within the human body (see Fig. 31.4).

The radiation exposure to the public resulting from normal nuclear operations is limited by federal regulation of engineering design and management controls to less than 5 percent of the natural background for the public. However, there is an overriding requirement that limits the radiation exposure to "as low as achievable." Radiation containment has been so successful that the federal limit is now being reduced from less than 5 percent to less than 1 percent.

The annual radiation exposure of nuclear workers is limited to 5 rem, and sites for nuclear operations are selected so that in the event of the maximum credible accident the exposure of the public will be less than 25 rem. The levels of radiation exposure were first established on the basis of historical data, and since then they have been periodically evaluated by the International Commission on the basis of continuing experience.

In 1975 the U.S. Energy Research and Development Agency sponsored the Reactor Safety Study, the most thorough hazard analysis ever performed[3] for any technology. It concluded that on the basis of past licensing review practices, the hazard from existing light water type power reactors was orders of magnitude less than other commonly accepted

TOTAL NATURAL DOSES (mrad/year)✓

Man over granite		Man over sedimentary rock	Man over sea	Large fish in sea		Micro-organism in sea	
10.000	m.s.l.			at surf.	100m	at surf.	100m
207	142	75	52	64	30	39	5

✓ A mrad is one thousandth of a rad, the unit for indicating absorbed energy in tissues.

Fig. 31.4. Natural radiation. (*"The Effects of Atomic Radiation on Oceanography and Fisheries," NAS-NRC Pub. 551/1957.*)

hazards. It further concluded that human error would be the expected cause of any accidents that might occur. This study also found that the cost of the nuclear power reactor accidents would be several orders of magnitude less than the cost of other natural and human-caused events to which the public is accustomed (see Fig. 31.5).

In the 1950s the Nuclear Safeguards Program was established by the United Nations to administratively reduce the risk of nuclear proliferation. The objective of the Safeguards Program is to prevent the diversion of fissionable material, primarily plutonium, from the reprocessing of spent fuel. A force of on-site inspectors is employed by IAEA in this activity. This program is administered by the IAEA and is generally considered to be successful. Countries join this program voluntarily and do not include their weapons production facilities. India did not choose to join but developed its own nuclear bomb. It is believed that several nonmember countries have or are developing nuclear bombs.

In 1980 a study of proliferation control, titled "International Nuclear Fuel Cycle Evaluation," was carried out.[4] President Jimmy Carter of the United States had called for this study, in which 50 countries participated and which required several years to complete. The study concluded that only administrative controls such as those already in place would be effective, and these could only delay really determined nations from diverting nuclear fuel from power reactors to military weapons.

Fig. 31.5. Frequency of property damage due to natural and human-caused events. (Notes: (1) Property damage due to auto accidents not included. (2) Approximate uncertainties for nuclear events are estimated to be represented by factors of 1/5 and 2 on consequence magnitudes and by factors of 1/5 and 5 on probabilities. (3) For natural and human-caused occurrences the uncertainty in probability of the largest recorded consequence magnitude is estimated to be represented by factors of 1/20 and 5. Smaller magnitudes have less uncertainty.)

THE EARTH'S ENERGY SUPPLY AND USE

The earth is provided with two sources of energy, the capital resources deposited within the earth during its formation and the income resource continuously beamed onto the earth from the sun. Both are essential to human existence.

In the 1950s, global energy analyses were performed by King Hubbard for the U.S. Geologic Survey and by Palmer Putnam for the Atomic Energy Commission. Hubbard's work reported the income energy from the sun to be equivalent to 178,000 million megawatts (Fig. 31.6). At noon on a cloudless day a square meter of the earth's surface facing the sun receives approximately one kilowatt. The income resource also includes a very small contribution, less than 0.02 percent, from the gravitational forces of the sun, moon, and earth and from the thermal and nuclear sources within the earth. Part of the solar energy appears in the form of water and wind power, wood, alcohol, garbage, cow dung, solar heaters, and photoelectric generators.

Putnam's study, titled "Energy in the Future," was primarily concerned with the capital energy resources. These analyses are continued today by the Energy Information Administration in the U.S. Department of Energy (see Fig. 31.7).

Fig. 31.6. World energy flow sheet. (*Hubbert, U.S. Geological Survey.*)

The capital resources are the fossil and nuclear fuels present in the crust of the earth. The fossil fuels were created by the interaction of the sun with the terrestrial flora and fauna. This was a very low-yield process that over the five billion years the earth has existed captured about two days of the solar energy in the form of coal, oil, gas, and other combustibles. The recoverable quantities are difficult to estimate.[9,10,11,12]

The nuclear fuels were created in the cosmic event that created the universe and were deposited in the earth as it took form. There are two families of nuclear fuels, those for fission (uranium and thorium) and those for fusion (protium, deuterium, helium-3, and lithium). Only uranium fission has been developed as a commercial source of nuclear energy. Although fusion has been developed as a military weapon, the hydrogen bomb, it is premature to include the fusion fuels in the world's inventory of capital energy. The technology for controlled fusion is not available, nor is development of a controlled fusion process expected in the next several decades. When available it would increase the capital supply to a level greater than that from all other sources combined.[8]

In 1990, the National Academy of Science completed an energy study that concluded that commercial fusion power may be required by the year 2050. By that time 40 percent of the oil reserves will have been consumed, and current natural gas reserves can supply only about half of the projected needs. The discovery of additional deposits and improved production might meet the need for natural gas, but at a higher price.

Little use was made of the earth's capital energy until the nineteenth century, and now 5 to 8 percent of that resource may have been consumed. Yet even though most of the fossil fuels remain, there are increasing problems with their extraction and distribution and the pollution that is incidental to their use. The demand for energy is increasing as the world continues to industrialize. The annual per capita consumption in the United States is

Fig. 31.7. U.S. energy flow, 1988 (quadrillion Btu). (*Annual Energy Review, U.S. Department of Energy, 1988.*)

180 million Btu's versus 70 million in Russia, 10 million in Brazil, and 5 million in India. The increasing demand will stem from population growth and increases in the standard of living. It will be accompanied by increased energy use for the extraction of marginal resources and the treatment of wastes before their release into the earth's environment.

Petroleum is and will remain the major source of merchanted energy for the next century. However, petroleum production has passed its peak in the United States, and in several decades it may peak in the rest of the world. However, the recent NAS energy study projects a 30 percent increase in production by 2010.

Coal can supply U.S. needs for the next hundred years, but there is increasing concern regarding pollution from both coal and petroleum. Unfortunately for the developed nations, using pollution-free energy sources such as nuclear and solar will do comparatively little to reduce atmospheric pollution. The sources of such pollution are worldwide. About 6 billion tons of CO_2 was generated in 1985, and this pollution source is expected to increase to 30 billion tons/year by 2060. In the 1980s only 10 percent of the 6 billion tons/year came from the United States.

The benefits of nuclear power obtainable with the present technology depend on the availability of uranium resources. For example, using the present light water reactor technology, each 1000-megawatt reactor requires about 10,000 metric tons of uranium during its 40-year operation. Because of limited uranium ore reserves in the United States and the lower cost of uranium from other countries, about 50 percent of U.S. consumption currently is imported. However, U.S. reserves should increase significantly when exploration again becomes profitable. In addition, the energy extracted from the uranium can be significantly increased by reprocessing the spent fuel. If the recovered mixed uranium and plutonium were used in a heavy water reactor, there might be as much as a 68 percent increase in energy generated. A further increase by a factor of ten might be achieved by using a thorium cycle or going to a liquid metal cooled reactor. However, these advanced technologies will result in higher electricity cost. In view of the current low cost for natural uranium, storage rather than reprocessing of spent reactor fuel creates a beneficial fission fuel reserve.

NUCLEAR PROCESSES

The science of the nuclear processes has a long history. Democritus in the fifth century B.C. started the search for the elementary particle of nature he called the atom. John Dalton, about 1803, related atomic masses to hydrogen, and Mendeleev in 1869 correlated these masses with the atoms' chemical properties and hypothesized missing elements. Many thought this ended the adventure, but in 1895 Roentgen discovered X rays, and a year later Becquerel discovered natural radioactivity. A short time thereafter, in 1902, Rutherford and Soddy proposed that radioactivity was related to atomic changes, and in 1905 Einstein equated mass to energy. In 1932 Curie and Joliot discovered artificial radioactivity, and in 1938 Meitner and Frisch recognized that uranium fissioned when bombarded with neutrons. In 1942 Enrico Fermi and Glen Seaborg led development of the first controlled neutron fission reactor to produce plutonium-239.[13]

The nuclear processes of most interest to the nuclear industry are radioactive decay and the transmutation of nuclides. Whereas chemical processes relate to the interactions of orbital electrons of the atom, nuclear processes relate to interactions of neutrons, charged particles, and nuclides with the neutrons and protons in the nucleus of the atom. As noted above, there are now known about 3000 nuclides and isomers; and only 287 of these are naturally occurring. More continue to be found. As Mendeleev invented the chart of the chemical elements, Emilio Segré invented the chart of the nuclides to give order to the nuclear properties and processes.[14]

TABLE 31.4 Radioactive Decay Processes

Reference process equation:

$$^{238}_{92}\text{U} \xrightarrow[\text{half life}]{4.5\text{E}9\text{y}} {}^{230}_{90}\text{Th} + {}^{4}_{2}\text{He} + \text{gamma} + \text{energy}$$

Thorium-230 also is radioactive, and the decay process continues through ten more radioactive nuclides before reaching stable lead-10.

	Nuclide			
Decay Process	*Parent*	*Daughter*[(b)]	*Half-life*[(a)]	*Energy (MeV)*
Alpha (α, $^{4}_{2}$He)				
	Pu-238	U-234	88y	5.5
Beta (β^-, °e)				
	Cu-64[(c)]	Zn-64(s)	13h	0.57
Positron (β^+, °e)				
	Cu-64[(c)]	Ni-64(s)	13h	0.65
Orbital electron Capture (EC, °e)				
	Cu-64[(c)]	Ni-64(s)	13h	1.68
Internal Transition (IT)				
	Tc-99m	Tc-99	6h	0.14
Fission				
	U-235	FP[d] ,n	Prompt	190
	Cf-252	FP ,n	2.6y	6
Fusion				
	D/T	He-4(s), n	Prompt	17

Notes: (a) Decay rate calculated by $T/2 = \ln 2(t - t_0)/\ln (N/N_0)$, where t = time and N = disintegrations/sec. (b) Daughters always include gamma rays. (c) Cu-64 has three decay processes (β^- 38%, β^+ 19%, and E, C 43%). (d) Fission products. (s) stable nuclides; others are radioactive.

Radioactive Decay

Radioactive decay occurs when an unstable atomic nucleus emits particles and/or energy to achieve a stable state. The nucleus contains neutron and protons along with the energy that binds them together. Many reactions are involved in radioactive decay (see Table 31.4). These reactions are characterized by the type, energy, and rate of radiation emitted. A different nuclide results from the decay process, which may also be unstable. This nuclide then will decay with its own unique radiation. Naturally occurring uranium-238 passes through 12 nuclides before reaching a stable end product, lead-210.

The types of radiation include:

- Electrons, called beta particles when negatively charged and positrons when positively charged.
- Helium ions, called alpha particles.
- Electromagnetic energy, called gamma or X rays, which accompanies all other emissions.
- Hydrogen ions, called protons.
- Deuterium ions, called deuterons.
- Neutrons.
- Nuclide ions from fission and fusion.

The energy of these emissions covers a wide range of values but is typically 190 million electron volts (MeV) for fission, 17 MeV for fusion, 5 MeV for alphas, 1 MeV for gammas, and 0.5 MeV for betas.

The rate of radioactive decay is expressed as the half-life, the time required for half of the unstable nuclide to decay to the next lower energy state. The half-lives range from less than microseconds to greater than a trillion years.

As radiation is not detected by the human senses, special detection and measurement methods are necessary. The silver halides in photographic film are sensitive to radiation as they are to the electromagnetic energy of visible light; in fact, it was photographic film detection that led to the accidental discovery of radiation. This technique continues to be used, particularly for the study of cosmic radiation. A closely related method is the use of crystals, for example, sodium iodide and germanium silicide, that emit light when subjected to radiation. This light is measured with photoelectric cells that can be tuned to measure the strength of the radiation. Another popular method measures the discharge of electric condensers in the form of gas-filled ion chambers. The radiation passing through the chamber ionizes the gas and discharges the condenser. Ion chambers are particularly useful for soft alpha and beta radiation measurement. The alpha emitter may be placed within the ion chamber. The Geiger counter is an ion chamber with its spontaneous discharge hooked to an audio speaker, whose resulting clicking provides an ominous background noise. All these detection methods have special uses but at the same time a wide range of applications. For each application they are calibrated with radioactive standards that are traceable to national and international reference standards.

Fission

Fission is a relatively simple process. Neutrons at room temperature are captured by a fissionable material, producing an unstable nuclide, part of which promptly fissions, yielding energy and fission products. The fission products generally include more than two neutrons, and under certain circumstances a fission chain reaction can be sustained. Under special conditions where more than one neutron per reaction is captured by the fissionable material, a violent explosion results. However, considerable effort is required to hold the material together long enough to obtain a high-yield explosion. In a controlled nuclear fission process, such as in a nuclear power reactor, only one of the neutrons is captured by other fissionable material. The remaining neutrons are used to produce more fissionable material, are absorbed in other components, or leak from the reactor.

Uranium-235. Natural uranium contains 0.711 percent uranium-235. The nuclear industry is still completely dependent on this single nuclide, which is the only naturally occurring one that can be effectively fissioned with thermal (log energy) neutrons. A controlled nuclear fission chain reaction was first achieved on December 2, 1942 under the west stands of the football stadium at the University of Chicago. The Chicago event was achieved in an atomic pile, a stacked array of natural uranium bodies in channels of a large graphite block. The graphite was present to reduce the energy of the neutrons to a lower level where they are most readily captured in U-235. (See Table 31.5.) Most of the fission energy, 180 million electron volts (Mev), is in the kinetic energy of the fission nuclides and is instantly captured in the uranium. The remaining energy, 20 Mev, is released in the form of radiation during decay of the radioactive fission products, strontium-90 and cesium-137, have lives of 30 years; and when this radiation is released, it is converted into thermal energy.

The fission neutrons at birth have energies of 2 Mev, which in a thermal reactor are reduced to thermal level, about 20 eV, for more efficient capture. There are many other materials in the reactor core also competing for the neutrons, including: the moderator (the material used to slow down or thermalize the neutrons); fertile nuclides that produce additional fissionable material (discussed in a later section); neutron poisons present in control rods; the coolant; and fuel element cladding.

Neutron poisons, which are nuclides with very high cross sections for the capture of neutrons, are an essential part of the nuclear reactor core. By their addition and withdrawal the nuclear chain reaction is controlled with a multiplication factor of one, that is, one fission per fission in the previous generation.

TABLE 31.5 The Fission Process

Reference equation:

$$^{235}_{92}U + ^{1}_{0}n \longrightarrow ^{236}_{92}U \xrightarrow{84\%} \text{Fission products} + 190\ \text{MeV}$$

$$^{236}_{92}U \xrightarrow[2E7y]{16\%} ^{232}_{90}Th + ^{4}_{2}He + 4.5\ \text{MeV}$$

Fission Cross Sections[a] *vs. Neutron Energy:*

Energy (keV)	*Thermal*[b]	*10*	*100*	*1000*
U-235	580	3.3	1.7	1.2
Pu-239	750	1.9	1.8	1.7
U-233	530	4.5	2.5	2.0
Neutron Yield/Capture vs. Neutron Energy:				
U-235	2.1	1.8	1.9	2.3
Pu-239	2.1	2.0	2.4	3.0
U-233	2.3	2.2	2.3	2.4

Thermal Neutron Capture Cross Sections:

Fissionable Nuclides		*Moderating Materials*	
U-235	690	H_2O	0.33
Pu-239	1030	D_2O	0.0006
U-233	580	Helium	0.007
Fertile Nuclides[c]		Carbon	0.004
U-238	2.7	Sodium	0.01[d]
Th-232		*Neutron Poisons*	
Fuel Cladding		Boron	760
Zircaloy	0.2	Cadmium	3,300
Stainless steel	3	Gadolinium	46,000

(a) Cross section in barns (10^{-24} cm^2). (b) Thermal energy at 20°C, 2200 m/sec, 0.0253 eV. (c) Capture neutrons to produce fissionable nuclides. (d) At 100–1000 keV.

To achieve high fuel burnup (utilization) in a light water reactor, boron (a poison) is present in the coolant during the early part of the cycle and then is removed as the fuel is consumed and poisons from the fission products accumulate. Cadmium, a lesser poison, is placed in control rods for fine tuning. Gadolinium, with a very high cross section for neutron absorption, finds use in emergency shutdown of heavy water reactors. Table 31.5 shows cross sections and neutron yields as a function of neutron energy for the principal fissionable materials.

Plutonium-239. Plutonium-239 represents a fruitious phenomenon. Whereas U-235 is the only significant fissionable nuclide in nature, its major isotope, U-238, captures a neutron to produce another fissionable nuclide plutonium-239. It is a more efficient material than U-235 both for the nuclear breeding cycle and for use as a nuclear explosive because its neutron yield is highest at high neutron energy.

Uranium-233. A second fissionable isotope, uranium-233, can be produced from naturally occurring thorium. It does not present an economically attractive option at present because of its dependence on highly enriched U-235 to bring the thorium cycle into operation and the large R&D expenditures required to develop the technology.

Fusion

At present the only application for fusion is in thermonuclear military explosives, where the necessary temperature and pressure are achieved with a plutonium fission initiator. The search for the science and technology needed to make fusion a viable source of energy for commercial electric power generation is the basis for one of the world's most important and challenging R&D efforts. The development of a controlled nuclear fusion reactor could provide a virtually unlimited source of energy. Like fission, fusion eliminates the atmospheric pollution associated with the use of fossil fuels. Also, it could significantly reduce the waste disposal problem of energy-producing processes.[15,16]

In the Tokamak fusion reactor depicted in Figure 31.8, electric current to the poloidal coils on the primary magnetic transformer generates the axial current in the secondary plasma composed of deuterium and tritium ions. These ions are heated to ignition temperature and then the reaction becomes self-sustaining. The toroidal field coil suspends the plasma away from the metal conducting walls. Contact with the wall would both cool the plasma below ignition temperature and contaminate the plasma with heavy ions. The relevant reactions are given below.

Deuterium-Tritium Fusion

1. $^{2}_{1}H + ^{3}_{1}H \longrightarrow ^{4}_{2}He + ^{1}_{0}n + 17.6\ \text{MeV}$

Fig. 31.8. Tokamak fusion reactor. (*Courtesy U.S. Department of Energy.*)

Tritium production in a fusion reactor blanket (or a fission reactor as is now done to produce tritium for the weapons program).

2. $^{6}_{3}\text{Li} + ^{1}_{0}\text{n} \longrightarrow ^{3}_{1}\text{H} + ^{4}_{2}\text{He}$

3. $^{7}_{3}\text{Li} + ^{1}_{0}\text{n} \longrightarrow ^{3}_{1}\text{H} + ^{3}_{2}\text{He} + ^{1}_{0}\text{n}$

Protium Lithium fusion would produce charged particles (90 percent of the energy in helium ions) for direct conversion to electricity, but higher temperatures and pressure would need to be achieved.

4. $^{1}_{1}\text{H} + ^{6}_{3}\text{Li} \longrightarrow ^{3}_{2}\text{He} + ^{4}_{2}\text{He} + 4\,\text{MeV}$

5. $^{1}_{1}\text{H} + ^{7}_{3}\text{Li} \longrightarrow ^{4}_{2}\text{He} + ^{4}_{2}\text{He} + 17\,\text{MeV}$

Of the several fusion reactions, deuterium–tritium fusion is the most feasible, as it has the lowest ignition temperature, 4E7 °K.* (See Reaction 1.) Deuterium comprises 0.15 percent of naturally occurring hydrogen, whereas tritium is produced by neutron fission of lithium-6 that is irradiated in a blanket surrounding a nuclear reactor core. Nuclide separation is required to produce the deuterium and possibly the lithium. The tritium would be produced in a surrounding lithium blanket of the fusion reactor to sustain the process. (See Reactions 2 and 3.) The tritium for hydrogen bombs is produced in nuclear fission reactors using these processes. The alternative fusion reactions use protium and lithium yield charged helium ions. (See Reactions 4 and 5.) These would contain 90 percent of the fusion energy and could be directly converted to electricity, but higher temperatures and pressures would be required.

Within the limits of our present knowledge, a very high temperature, about 40 million °C, at high pressure for a finite period of time, is required to initiate the fusion reaction. The fusion temperature in the sun is only 15 million °C, but the high solar pressures cannot be achieved on earth.

There are now two approaches to developing a fusion reactor, magnetic containment and inertial confinement. The major effort, magnetic containment, uses the Tokamak concept to compress and heat the reactants in a plasma isolated in a vacuum away from the walls of the reactor. The fuel is injected as pellets and heated with electric ohmic energy to strive for the ignition temperature. Research along this line started in the early 1950s, and in 1968 the Russians discovered the Tokamak principles. (Tokamak is the Russian acronym for "toroidal chamber with magnetic coil.") Massive amounts of electrical energy are required to heat and compress the plasma. Ohmic energy now is supplemented with induction, microwave, and neutral beams techniques. The time of heating has been extended from microseconds to over a minute using these additional methods.

There are two immediate objectives in fusion research, first to achieve ignition and then to achieve parity (produce more energy than is required to maintain steady state operation). This research represents a large international activity involving principally Germany, Japan, Russia, and the United States, which collectively are spending more than a billion dollars a year. Although there is continuing progress, another decade may be required to achieve controlled ignition, and this will need to be followed by the effort to

*4E7 is the exponential statement of 4×10^7.

TABLE 31.6 Nuclide Production Process

$$A = N\sigma\phi(1 - e^{-\lambda t})$$

where:

$(1 - e^{-\lambda t})$ = saturation factor: for small values of λ, it equals λt; for $\lambda = 1$, it equals 0.5, the optimum irradiation time.
A = disintegrations per second; curies = 3.7E10 d/s.
N = number of atoms; Avogrado's number = 6.023E23 atoms/gram atomic wt.
σ = sorption cross section (1E − 24) (may be written 10^{-24})cm^2, called a barn.
ϕ = irradiation flux, particles/sec-cm; for electromagnetic accelerators reported in microamperes (1 μa = 2.3E11 particles)
λ = decay constant, natural logarithm 2/half-life.
t = irradiation time, seconds.

Examples:

1. Co-60 production (Co-59 target at 1E14 neutron flux, $\sigma = 37$):

Irradiation Time (y)	*Curies Co-60/gram target**
1.1	130
2.2	225
5.5	370

2. Pu-239 production (U-238 target at 1E14 neutron flux, $\sigma = 2.7$):

	Grams Pu-239/metric ton target
1	4,800
2	9,600
3	14,400

*These quantities ignore depletion of the products through decay and conversion to other products.

achieve parity. When ignition and parity are fianlly achieved and understood, spending will need to be substantially increased to establish commercial feasibility. It is now realized that electricity from fusion will be more costly than that from fossil and fission processes. It is only as these resources are depleted that fusion will become economically feasible.

Nuclide Production

Nuclides are produced by capturing a nuclear particle in the nucleus of the target atom, which thus is transmuted to a different atom, one that in most cases is unstable and decays, as discussed in an earlier section.[17]

The target atom may be any stable or radioactive atom. It is bombarded, and the process generally is called irradiation, with production of many of the nuclear particles discussed in the previous section. Most often the particle is a neutron produced in a nuclear reactor, but frequently it is a charged particle raised to the required energy level in an electromagnetic accelerator. The charged particles often are protons but may be deuterons, alpha particles (He-4 or He-3), or heavier ions.[20]

The rate of nuclide production is dependent on the number and the energy of particles bombarding the target, the cross section of the target and the half-life of the product (see Table 31.6). The decay of these products by fission and fusion is the source of tremendous amounts of energy, as already discussed.

Fission Products

The fission process produces radioactive as well as stable nuclides with masses ranging from 72 to 167 and with two broad peaks in the regions of 95 and 138. The masses are identified rather than the specific nuclides because in fission many short-lived nuclides

TABLE 31.7 Fission Products

Mass	*Long-lived Nuclides*	*Percent Yield*[a] *U-235*	*Pu-239*
72	Zn	0.001	0.001
79	Se	0.044	0.044
85	Kr	1.3	0.58
90[b]	Sr	5.8	2.0
99	Mo	6.1	6.2
102	Pd(s)	7.7[c]	28.1
103	Rh(s)	3.0	7
124	Xe(s)	23.7	27.2
131	I	2.9	3.9
133	I, Xe	6.7	7.0
137	Cs	6.7	6.6
140	Ba, La	6.2	5.4
167	Dy	0.001	0.001

[a] Because two fission products are emitted for every binary fission, the cumulative yield for the mass chains is 200 percent.
[b]Mass-90 beta decay chain (half-life): Br(2s), Kr(32s), Rb(4m), Sr(29y), Y(2.6d), Zn(stable).
[c]This is equivalent to 1.2 kg of palladium/metric ton of spent fuel burned to 32,000 MWd. One megawatt-day of thermal energy from fission is approximately equivalent to one gram of fission products.
[s]Stable products of the mass chains. Multiple chains contribute to the high yields for palladium and xenon.

are produced that quickly decay by beta emission to a long-lived or stable member of the mass chain.[14]

The yields of selected mass chains that result from fission of U-235 and Pu-239 are shown in Table 31.7. In addition to the fission products with masses roughly half that of U-235, neutrons, tritium, helium, and beryllium are products of triple fission. There are significant differences in some of the yields from U-235 and Pu-239.

Although the fision products could be recovered as by-products from the waste from spent nuclear reactor fuel, special neutron irradiation of highly enriched uranium (isotopically separated uranium-235) followed by chemical separation is the normal production method. The major products, molybdenum-99 and iodine-131 with fission yields of 6.1 and 6.7 percent, respectively, have important medical applications. Mo-99, with a half-life of 2.7 days, is particularly interesting because its short-lived decay product, technetium-99m with a half-life of 6 hours, is the product actually used. To make this short-lived product available, the Mo-99 is sorbed in a silica gel column, and the Tc-99m is eluted as needed.

During the 1980s, large-scale processing of the defense wastes at Hanford extracted and encapsulated 100 megacuries of strontium-90 and cesium-137 to reduce heat generation in the waste storage tanks. Krypton-85 has been recovered from the airborne effluents at the Idaho processing plant.

In addition, development studies have addressed the recovery of stable rhodium, xenon, technetium, and palladium because of their limited availability in nature. For palladium this would amount to about 40 kg per year from a 1000 MWe reactor. However, a small amount of palladium-107 (half-life 6E6y) is present that could complicate its use. This is also a problem for technetium, which is 100 percent technetium-99 with a half-life of 2E5y.

The mass-133 chain represents a fission reactor problem. Iodine-133 with a 21-hour half-life decays to 5-day xenon-133, which has a high cross section for thermal neutrons. This was discovered when the first Hanford reactors unexpectedly shut down after a few hours' operation. Fortunately, these reactors were built with provision for the extra fuel that could override the xenon effect. However, when a reactor is shut down, the xenon-133 grows in from its I-133 parent. Unless it is started up again in a few hours, several days must pass before the Xe-133 has decayed enough to allow restart.

Neutron Transmutation Products

The production of neutron transmutation products involves the neutron irradiation of target materials followed by chemical separation. Although this promises high purity with high specific activity, the isotopic composition of the target along with other competing reactions frequently yields a spectrum of isotopes of the desired nuclide (see Table 31.8).

TABLE 31.8 Neutron Products

Transmutation:

$$^{238}_{92}U + ^{1}_{0}n \xrightarrow[(2.7)]{} ^{239}_{92}U - ^{0}_{1}e \xrightarrow{24m} ^{239}_{93}Np - ^{0}_{1}e \xrightarrow{2.4d} ^{239}_{94}Pu$$

$$^{239}_{94}Pu + ^{1}_{0}n \xrightarrow[(270)]{} ^{240}_{94}Pu + ^{1}_{0}n \xrightarrow[(290)]{} ^{241}_{94}Pu + ^{1}_{0}n \xrightarrow[(400)]{} ^{292}_{94}Pu$$

Others:

K-43/2.2h	I-131/8d	Gd-153/242d
Cu-67/62h	Po-210/138d	H-3/125y
P-32/14d	C-14/5730y	

Activation:

$$^{59}_{27}Co + ^{1}_{0}n \xrightarrow[(37)]{} ^{60}_{27}Co - ^{0}_{1}e \xrightarrow{5.3y} ^{60}_{28}Ni$$

Others:

Mo-99/2.8d	Ir-192/74d	Ni-63/125y
Cr-51	Sn-113/118d	Cl-36/3E5y
Fe-59/45d	Tm-170/130d	

Note: Figure in ( ) is thermal neutron cross section.

Plutonium-239 and tritium for use as military explosives are the two major transmutation products. The nuclear process for Pu-239 production is the same as for energy generation, but there are some differences: (a) metallic natural uranium clad with aluminum facilitates later dissolution for plutonium recovery, and so the reactor operates at relatively low temperature; (b) the irradiation cycle is limited to a few months to minimize the PU-239 conversion to Pu-240 and Pu-241; and (c) a carbon or a heavy water moderator is used to increase the neutron efficiency.

More complex is the production of Pu-238 and californium-252. For Pu-238 there are two irradiation cycles: the first starts with natural uranium but maximizes the neptunium-237 yield; then the Np-237 is separated, converted into aluminum-clad oxide targets, and irradiated to Pu-238 in a second cycle. The total process requires double neutron capture.

Californium-252 production was especially challenging, as it involved the sequential capture of 14 neutrons along with the intermediate separation and fabrication of two intermediate targets (americium and curium isotopes) when starting with Pu-239.[18] This production campaign lasted ten years, produced about 10 g of Cf-252, and then was terminated. The product was evaluated as a neutron source but had insufficient value to justify continuing production.

Neutron Activation Products

The activation production cycle requires only neutron irradiation without chemical separation. The target and the product are the same chemical element but have different nuclide compositions. The specific radioactivity of the product is a function of the nuclide composition of the target, the neutron flux environment, the irradiation time, and the half-life of the product nuclide, along with the nuclear cross sections of target and product nuclides (Table 31.8).

The major neutron activation product is cobalt-60 with a reasonably long half-life (5.3 years). It emits high-energy beta particles and gamma rays appropriate for radiographic, irradiation, and isotopic power applications. It is an ideal product because naturally occurring cobalt is uniquely mono-isotopic cobalt-59 as well as a relatively stable metal.

TABLE 31.9 Nuclide Production by Charged Particles

Typical reaction:

$$^{14}_{7}N + ^{1}_{1}H \longrightarrow ^{15}_{8}O - ^{4}_{2}He \xrightarrow{*} ^{11}_{6}C - ^{0}_{1}e \xrightarrow{20m} ^{4}_{5}B\,(\text{stable})$$

This reaction normally is reported as N-14(p,α)C-11. Natural nitrogen gas is bombarded in a cyclotron with a 30-microampere current of 4 to 13 MeV protons (p). O-15 is produced, but an alpha particle is promptly (*) ejected, producing C-11, the desired product. The yield for a 20-minute bombardment is 0.2Ci.

$$^{58}_{28}N + ^{1}_{1}H \longrightarrow ^{59}_{29}Cu - 2^{1}_{1}H \xrightarrow{*} ^{57}_{27}Co - ^{0}_{1}e \xrightarrow{270d} ^{57}_{26}Fe\,(\text{stable})$$

Ni-58[p,2p]Co-57. Enriched Ni-58 (enriched in the Oak Ridge Calutron) is plated on a copper heat exchanger, bombarded for 5 hours with 1000 microamperes of 20 MeV protons. Cu-59 is produced but 2 protons are promptly ejected producing Co-57, the desired product. The yield for a 14 hr bombardment is approximately 28 microcuries per microampere hour.

Other (p,n) Processes: Li-7 to Be-7; Fe-56 to Co-56; Cu-65 to Zn-65; Ag-109 to Cd-109.

Other Processes and Products: N-14(d,n)O-15; O-16(p,α)N-13; Ne-20(d,α)F-18.

Furthermore the target nuclear cross section is good, and the product cross section is lower than the target's by a factor of ten. Another factor in cobalt-60's long list of advantages is the Canadian CANDU family of nuclear power reactors. They are heavy water–moderated and readily adapted to Co-60 production, and 16 are committed to that use. In 1990, 50 megacuries was produced with a catalog price of $1.20/curie.

Other important activation products include molybdenum-99 and iridium-192. However, the Mo-99 is better obtained from fission, and already has been discussed in that section. The 74-day Ir-192 has a gamma with less energy than Co-60 and is used for radiography of less dense materials.

Charged Particle Transmutation Products

Many radionuclides can be produced in cyclotrons, thus avoiding the use of more costly nuclear reactors. Many research hospitals now have cyclotrons to provide the short-lived radionuclides of carbon, nitrogen, oxygen, and fluorine. The longer-lived products are produced commercially or in government laboratories.[19,20]

These radionuclides are produced by irradiating targets with beams of hydrogen ions (protons), but frequently deuterium ions (deuterons) are used (see Table 31.9). Some products require helium-4 and helium-3 ions. The typical process involves the capture of the proton with the prompt emission of a neutron. This is called a p,n reaction. However, in other cases there may be protons, alpha particles, or up to five neutrons emitted. The resulting products decay generally by positively charged electron (positron) emission, but also decay by capture of an orbital electron.

To produce the beam of high energy charged particles, a stream of hydrogen or helium is electrically ionized and accelerated in an oscillating magnetic field that holds the particles in a spiral path within the cyclotron. At the desired energy the particle beam is magnetically deflected to the target. The beam strength can vary from 4 to 800 MeV but is usually less than 30 MeV. The beam current ranges from 3 to 2500 microamperes. The target is designed to accommodate the ten to several hundred watts of thermal energy deposited by the beam. Gas and liquid targets are used for the short-lived products to accelerate processing for use. For the long-lived products that require longer irradiation and higher energy flux (up to a kilowatt per square centimeter), the target material may be plated on a water-cooled heat exchanger. The optimum irradiation time is equivalent

to the half-life of the product. The target materials in many cases are separated stable nuclides.

The Department of Energy's laboratories have accelerators that provide up to 800 MeV proton beams. Other machines support high energy physics research but occasionally are used to produce special nuclides such as xenon-127.

REACTOR MATERIALS PROCESSING

Reactor materials processing is concerned with preparation of the special materials used in nuclear reactors. This includes the nuclear reactor fuels along with the reactor vessels, plumbing, heat exchangers, coolants, and moderators. In this chapter only the nuclear fuel cycle is discussed. All the materials that go into the nuclear reactors are deliberately and thoroughly controlled by standards established by the nuclear industry with the participation of the Nuclear Regulatory Commission and administered by the American National Standards Institute. These materials include the uranium oxide pellets, zirconium cladding, and steel fittings of the fuel assemblies, the corrosion- and radiation-resistant steel used for the reactor vessels, and the various corrosion- and erosion-resistant steels for piping, valves, pumps, and heat exchangers. In addition there are the coolants (water, sodium and helium), the moderators (deuterium and graphite), and the neutron sorbers (boron, cadium, hafnium, and gadolinium). Each of these materials requires intensive processing to achieve safe and reliable operation of the nuclear reactors.

In this chapter only the isotope-enriched nuclear fuel cycle is considered.

Isotope Enrichment

Isotope enrichment is a major activity of the nuclear industry. The naturally occurring chemical elements are composed of isotopes with varying nuclear properties. By enriching the chemical element in the nuclide with the desirable nuclear property, the performance of the nuclear processes is enhanced. In addition, the sensitivity of isotopic analysis permits the use of enriched chemical elements in place of radioisotopes as tracers in research, medicine, and industry.

The separation of chemical isotopes is based on small differences in their physical and chemical properties. For the lower-mass isotopes, chemical exchange, distillation, and electrolysis have been used. For the higher-mass isotopes, techniques based on mass have been used, including gaseous diffusion, centrifugation, thermal diffusion, and ion activation.[21,22]

The most universal of these techniques uses the Calutron, an electromagnetic accelerator employing the first arc of a cyclotron (Fig. 31.9). In 1945 it provided the enriched U-235 that was used in an early atomic bomb. Since then it has been used to enrich small quantities of most of the stable isotopes. The separation is achieved by ionizing the feed material and then accelerating the ions through the field of a bending magnet. The differences in mass cause the heavier elements to pass through a larger arc with the ions trapped in carbon pockets located 180° from the feed point. The accelerating voltage and magnetic field are controlled to optimize performance. Separation factors of 10 to 100 are obtained, but the yields are only 1 to 10 percent.

Large-scale production has been applied to enrich U-235, deuterium, and lithium-6. Gaseous diffusion is the principal technology for U-235 and is discussed further as part of the fuel cycle. Dual-temperature chemical exchange processes are used to enrich hydrogen and lithium isotopes.

Natural hydrogen contains 0.16 percent deuterium and is enriched to 10 to 20 percent deuterium in pairs of two hundred plate columns.[23] The chemical exchange system is hydrogen sulfide gas and water. One column operates at 27 °C and the second at 220 °C, with the enriched product removed from the bottom of the cold column that flows to the second set of columns. From the second set it goes to distillation and then to electrolysis to yield 99.8 percent deuterium. The specific enrichments at the crossovers

Fig. 31.9. Oak Ridge Calutron for separation of isotopes. The first method to achieve large-scale separation of U-235 and today the source of research quantities of most stable and a few radioactive nuclides. (*Courtesy USDOE.*)

between processes is controlled by energy considerations. The deuterium is produced primarily for use as moderator and coolant for heavy water nuclear reactors.

The chemical exchange system employed for lithium-6 enrichment is lithium amalgam and aqueous lithium hydroxide. It also employs paired dual-temperature columns.

Cryogenic distillation is used to enrich carbon-13, nitrogen-15, oxygen-17, and oxygen-18. Typical of these processes is carbon monoxide distillation, which has a C-13/C-12 separation factor of 1.008 between the vapor and the liquid. The initial 20-meter packed column is tapered from 2.5 cm at the boiler to 10 cm at the reflux condenser. This brings the carbon-13 from 1.1 to 12 percent. In a second column it is brought to 93 percent.

THE URANIUM FUEL CYCLE

The uranium fuel cycle for the uranium–plutonium system is a multicomponent system of chemical process operations that begins with mining uranium ore from the earth as

the starting material and ends with the waste products in above-ground or below-ground terminal repositories. Some radioactivity is released in the air and water discharged to the environment at concentrations and quantities below those specified by federal regulations. The topics to be considered in this section are mining, milling the ore, uranium fuel preparation, and reprocessing.[24,25] The following section discusses radioactive waste management.

A similar set of processes has been partially developed for the thorium–uranium system but will not be discussed here because it is not expected to be employed in the next several decades. The important feature of the thorium cycle is that it could be used to achieve breeding (to produce more fissionable material than is consumed) in thermal reactors, but nuclear as well as chemical factors have frustrated this development. The increasing cost of the natural uranium supply for the uranium/plutonium cycle may, several decades in the future, justify development of the thorium cycle.

It should be noted that breeders would not reduce the demand for uranium ore for many decades because several LWR and/or HWR converters (which produce fissionable material, but less than consumption) are required during the run-in of a breeder cycle to equilibrium. The doubling time of a breeder (the time required for the breeder to produce sufficient fissionable material to start up a second breeder reactor) might be a significant part of its operating life. Furthermore, natural uranium will be required for the thorium cycle, if it is used, and for startup of the fusion cycle. The tritium for the fusion cycle will be made in nuclear reactors, as it now is for nuclear weapons. The nuclear industry will always be dependent on a continuing supply of uranium from ore.

Mining

Uranium mines are primarily the open pit type, but there is significant production from deep mines as well as from solution mining. Sometimes uranium is produced as a by-product to mining operations for vanadium, phosphate, and gold.

Where uranium ore deposits occur in permeable aquifers with low-permeability geologic formations above and below them, the uranium is extracted by circulating a leaching solution. The leach solutions are sulfuric acid, ammonium carbonate, or sodium carbonate–bicarbonate along with air, peroxide, or sodium chlorate as the oxidizing agent. Holes are drilled to the bottom of the ore body on a grid pattern, 50 to 200 feet apart. The holes are cased with screened sections through the ore body. Varying combinations of injection and withdrawal of the leaching agents are applied to recover the uranium, which is recovered at a very low concentration, and is stripped from the leaching agent by ion exchange. The leachate is adjusted and recycled.

Milling

In milling, the uranium is leached from the ore, separated from major contaminants, and converted to yellow cake. The uranium ores milled in the United States generally contain 0.05 to 0.2 percent U_3O_8, with an average of 0.1 percent; and to save transport costs, the mills are located near the mines. There are two major processes, one based on a carbonate leachant and the other on sulfuric acid (See Fig. 31.10.)

Carbonate Process. In this process the ores are leached with hot sodium carbonate for 24 hours, with sparging with air to provide oxidation. The leachate is cooled in counter-current heat exchangers, heating the carbonate solution for the next batch. The carbonate leachate is filtered on rotary drums, and the uranium is precipitated with sodium hydroxide and filtered. The filtrate is converted back to carbonate by sparging with carbon dioxide, usually from a boiler flue gas, and recycled. The dried precipitate in the form of sodium diuranate is about 90 percent U_3O_8. Carbonate leaching is fairly specific for uranium, and the product is fairly low in contaminants.

Fig. 31.10. Uranium ore mill processes: (A) carbonate leach, caustic precipitation process; (B) acid leach, ion exchange and acid leach, solvent extraction processes; (C) acid leach, resin-in-pulp process.

Sulfuric Acid Process. In the sulfuric process, the pulverized arc is leached for about 8 hours, with sodium chlorate or manganese dioxide added for oxidation. The leachate passes through a series of thickeners countercurrent to the flow of the wash water. The acid leachate is not very specific for uranium and therefore goes through ion exchange or solvent extraction for purification. With the RIP (resin-in-pulp) process, after thickening to avoid the filtration step the slime-containing leachate goes into tanks where screen baskets of anion exchange resins, mechanically sloshed up and down, sorb the uranium. After loading, the uranium is stripped from the resin with an acidified chloride or nitrate solution. In some plants the leachate is filtered and passes through solvent extraction or fixed-bed ion exchange. The solvent is a hydrocarbon-diluted 2-diethyl-hexyl phosphoric acid applicable to the recovery of both vanadium and uranium when both are present in the ore. These product streams are precipitated with sodium hydroxide and dried. An additional cycle of ion exchange may be required to achieve the desired product quality.

Fuel Preparation

Uranium fuel preparation takes the yellow cake, sodium diuranate, from milling to plants where it is converted to either (a) aluminum-clad metal for the weapons plutonium production reactors (and in the future for the liquid metal cooled reactors) or (b) to Zircaloy-clad UO_2 for electricity production in the light and heavy water power reactors (see Fig. 31.11).

For the production of these clad fuels, there are two general processing routes. The primary

Fig. 31.11. Uranium feed materials flow sheet.

Fig. 31.12. Uranium solvent extraction process for purification of ore concentrate and scrap; slurry feed eliminates clarification cost and losses. (*Courtesy USDOE.*)

route is through solvent extraction for purification, denitration to UO_3, and reduction with hydrogen to UO_2, followed by either hydrofluorination to UF_4 and reduction to metal, or fluorination to UF_6 followed by isotopic separation and reduction to UO_2. The other route starts with higher-grade yellow cake, skips solvent extraction, and goes directly to hydrofluorination, with distillation of the final hexafluoride added for purification. Isotopic separation and conversion to UO_2 complete the process. This discussion of fuel preparation also contains several paragraphs on uranium enrichment, gaseous diffusion, and uranium hexafluoride conversion to oxide to form fuel rods.

Solvent Extraction. The yellow cake is dissolved in nitric acid and extracted from this aqueous phase by 5 percent tributyl phosphate (TBP) in a hydrocarbon diluent. The diluent reduces the density and viscosity of the TBP, enhancing the aqueous/solvent phase separation. The extraction is very specific for uranium, with separation factors of 10^3 to 10^5. The product thus obtained is an aqueous uranyl nitrate solution (Fig. 31.12).

Denitration. The uranyl nitrate solution from solvent extraction is converted to UO_3 by evaporating the solution to a final boiling point of 120 to 140 °C, followed by calcination at 620 °C. The product characteristics are dependent on the type of calciner (pot, trough, or fluid bed) and significantly affect the subsequent steps (see Fig. 31.13).

Conversion to Uranium Hexafluoride. The calcined UO_3 is reduced to UO_2 with hydrogen, then converted first to UF_4 with hydrogen fluoride and finally to UF_6 using fluorine. These steps usually are carried out in fluidized-bed reactors, but the conversion

Fig. 31.13. Uranium denitration process prepares UO_3 for conversion to metal and UF_6. A continuous fluid-bed process has been developed. (*Courtesy USDOE.*)

Fig. 31.14. Fluid-bed system for UO_2 conversion to UF_4. (*Metropolis process.*)

Fig. 31.15. Uranium metal reduction process; a similar process is used for plutonium metal production. (*Courtesy USDOE.*)

to UF_4 sometimes is done in a stirred trough reactor.[27]

The fluid-bed process at the Allied Metropolis Plant uses a series of three fluidized beds (see Fig. 31.14). The first bed is fluidized with a 0.5 ft/sec flow of hydrogen and reduces the UO_3 to granular UO_2. Then it is hydrofluorinated to a granular UF_4 in a two-stage fluid bed to achieve efficient consumption of the hydrofluoric acid. Finally, in a fourth fluid-bed reactor the granular UF_4 burns in a fluorine atmosphere to UF_6 vapor. The fluorine is fed in through the ash at the bottom of the reactor to maximize the uranium yield, and then exits through the UF_4 screw feeder to minimize the loss of the high-cost fluorine.

UF_6 Purification. Distillation of the UF_6 is required for purification when the yellow cake does not go through solvent extraction. At Metropolis, the molybdenum, vanadium, and other impurities are removed in a pair of columns, one a 120-foot, 100-plate column operating at 200 °F and 85 psia to remove the high volatiles and the second a 45-plate column operating at 240 °F and 95 psia to remove the low volatiles. The UF_6 is condensed into 10-ton cylinders and delivered to the enrichment plant.

Reduction to Metal. Uranium metal is produced by bomb reduction of UF_4 with magnesium metal. In the process, granular UF_4 is blended with magnesium metal pellets and tamped into a steel reactor lined with the reaction by-product, magnesium fluoride (see Fig. 31.15). After the container is capped, it is placed in a furnace where the temperature is raised to the ignition temperature, at which the magnesium and UF_4 react spontaneously, and the reaction mass reaches a temperature sufficiently high for the liquid uranium metal to form a puddle within the bomb. This is a very empirical technology, dependent on the control of many physical factors. Failure will produce dispersed uranium shot instead of the desired regulus.

The resulting 350-lb uranium regulus, called the derby, is broken out, remelted in a vacuum furnace, and held at 1454 °C to volatilize and remove the impurities. It then is recast in graphite molds to produce the ingot. This is formed into 1- to 2-inch-diameter rods by

extrusion and rolling-mill operations followed by machining and cladding with aluminum. Its primary use is for plutonium production; however, some of the depleted metal is used for shipping-cask shields, military projectiles, and counterweights. The magnesium fluoride by-product is ground and screened to provide material for lining the metal reduction bombs.

Uranium Enrichment. Enrichment of uranium-235, from 0.711 percent as present in natural uranium, is essential to the economical operation of the light water reactors. The fuel life for these reactors is a function of the enrichment. With 3.4 percent U-235 fuel, the pressurized light water reactors produce 33,000 megawatt-days of energy/metric ton during the three years the fuel is in the reactors. In the naval submarine and ship reactors using highly enriched U-235, the fuel life exceeds ten years.

During the original development, gaseous diffusion was selected over electromagnetic separation and thermal diffusion to separate uranium isotopes. Enrichment represents about 4 percent of the cost of nuclear generated electricity. Although gaseous diffusion is still the major technology, three other technologies are now in various stages of development that might reduce this cost by a factor of four; they are based on gas centrifuges, jet nozzles, and lasers. The gas centrifuge and jet nozzle plants now are operating (see Table 31.10). Separation in the ultra-high-speed centrifuge depends on isotopic mass difference, as does the jet nozzle.[29] The last-named device is a pneumatic cyclone with a high-velocity gas flow induced by a high differential pressure.

TABLE 31.10 Enrichment Plant Capacities (million SWU/year)

	Gaseous Diffusion	*Gas Centrifuges*	*Jet Nozzle*
USA	19		
Europe	11	3.2	
Russia	10		
Japan		1.5	
UK	0.6	0.9	
China	0.2		
Germany			0.05
Brazil			0.01

Laser-induced separation is being developed, based on differential activation energies of the uranium hexafluorides. This latest enrichment technology uses laser irradiation of the hexafluoride vapor to selectively decompose the U-235F_6 and precipitate U-235F_5. This process promises to recover much of the U-235 that remains in the tails from gaseous diffusion plants. Some U.S. utilities are considering constructing a plant to use the laser technology.

Gaseous Diffusion. In the gaseous diffusion process, the UF_6 flows through a porous nickel membrane called the barrier. The heavier U-238F_6 flows more slowly than the U-235F_6, and the theoretical separation factor for an equilibrium stage is:

$$\alpha = \sqrt{\frac{^{238}UF_6(352)}{^{235}UF_6(349)}} = 1.0043$$

The optimum conditions require elevated temperature and reduced pressure with a pressure differential across the barrier sufficient to transport half the UF_6 flowing through that stage. However, by varying this flow ratio the stages are reduced in size as the enrichment of the U-235 is increased. This practice, called tapering, is very important because a large number of stages are required. In the largest stages, axial flow compressors driven by electric motors rated to 3300 horsepower transport the UF_6, and 640,000 kg of this material is circulated to produce one separation work unit (SWU). When it was operating at full capacity, about 150 metric tons of UF_6 was fed to the plant daily in the process described here. Tapering reduces the time required for the plant to come to equilibrium.

The Department of Energy's gaseous diffusion plant in 1980 had 10,812 stages, consumed 6000 MW of electric power and 1350 million gallons of water a day, and took months to come to equilibrium. The building housing the plant had a combined floor area of a square mile. Plans to further expand the

plant using gas centrifuges were scrapped in the 1980s when the expected growth of nuclear power was not realized for power reactors. In addition, other countries, including Russia, now offer such services.[28]

The capacity of processes employing diffusion and mass separation is reported on the basis of the work required, and is expressed as separation work units. (The SWU is a measure of the work required to separate uranium of a given U-235 content into two components, one having a higher and the other a lower U-235 content.) It takes 5.3 SWU to enrich natural uranium to 3.44 percent U-235. The Department of Energy facility has the capacity of 19 million SWU, and there is a world capacity of 46 million SWU (Table 31.11.)

TABLE 31.11 Separation Work Unit Requirements (SWU/kg product)[a]

Percent U-235	*SWU*
0.5	0.17[b]
0.6	0.11[b]
0.711	0.00
1.0	0.4
2.6	3.4
3.4	5.3
4.0	6.5
5.0	8.9
20.0	45.7
98.0	270

[a]Tails 0.2 percent U-235.
[b]Added for depleted fuel. Reference cost $1.20/SWU-kg.

Uranium Hexafluoride Conversion to Oxide. The UF_6 is hydrolyzed in an aqueous ammonia solution to precipitate ammonium diuranate, which then is calcined in a hydrogen atmosphere to UO_2. This is then granulated, pelletized, and fired in a hydrogen furnace to achieve high density. (See Fig. 31.16.) The oxide pellets are then ground to size and made into high density pellets, which are loaded into zirconium tubing to form fuel rods. There are approximately 80 miles of these rods in a nuclear power reactor. It is expected now (1992) that after burnup in the reactors, these rods will be go to final disposal. However, the spent fuel contains about 4.5 percent fisionable material, U-235 and Pu-239.

Spent Fuel Reprocessing

The reprocessing of spent nuclear fuel is the most sensitive operation in the nuclear industry. It is there that the largest quantities of fissionable and radioactive nuclides are

Fig. 31.16. UF_6 conversion to UO_2 by ammonium diuranate precipitation. (*Courtesy USDOE.*)

Fig. 31.17. General solvent extraction process for spent fuel recovery.

handled in aqueous solution. These large-scale operations require both remote control and remote maintenance of the plant to protect the workers from radiation. In addition, the air and water effluents along with the solid refuse must be closely monitored to assure that the public is protected. Finally, the fissionable material requires strict accountability to ensure that it is not diverted to unauthorized uses.

As nuclear safety is a foremost public concern, it is well to note that the operation of the plutonium recovery plants since 1945 has met design specifications. There have been no significant health effects on either the workers or the public. The recovery plants were designed with large underground tanks for interim storage of the aqueous waste. The technology for final waste disposal is still unresolved and is discussed elsewhere in this chapter. Figure 31.17 shows a generalized flow diagram for nuclear fuel recovery via solvent extraction.[30]

The centerpiece of spent fuel reprocessing is the Purex process (Plutonium-Uranium-Extraction). The solvent is tributyl phosphate in a hydrocarbon diluent, and the process was first used at the Ames Laboratory for uranium purification, then at Oak Ridge National Laboratory for spent fuel. Although other processes were used in earlier days, the Purex process, with various modifications, is now used for many chemical separations in the uranium fuel cycle. It first was used on a large scale to recover uranium from the bismuth phosphate waste generated by the original plutonium recovery operation. This technology was shared with the world at the first Atoms for Peace Conference in Geneva in 1955.

A typical Purex process includes the following:

1. Aging the spent fuel to reduce by radioactive decay the 8-day iodine-131 that would be released to the atmosphere during dissolution.
2. Shearing the Zircaloy clad fuel rods into approximately one-inch lengths so the UO_2 can be dissolved.
3. Dissolving the UO_2 in nitric acid.
4. Solvent extraction to separate the plutonium, uranium, and fission products.
5. Anion exchange to clean up the plutonium.
6. Sorption to clean up the uranium.

Modifications to this process can be made to effect recovery of neptunium, americium, curium, californium, strontium, cesium, technetium, and other fission nuclides. The efficient production of specific transuranic

products requires consideration of the irradiation cycle in the reactor and separation of intermediate products for further irradiation.

The spent fuel as it leaves the typical light water cooled reactor contains 3 percent fission products and 0.6 percent transuranic products. The fuel provided about 32,000 megawatt days of thermal energy/metric ton of uranium. The fuel as charged to the reactor contained 32,000 g of fissionable material, U-235; and one gram of fissionable material produces about 1 megawatt day of thermal energy. Therefore, 30,000 g of fissionable material, or all the fissionable material charged, has been consumed. However, there is still about 44 percent of fissionable material in the spent fuel. This is present as 7600 g of the original U-235 along with 4800 g of Pu-239 and 1100 g of Pu-241. This means that 7600 g, or 25 percent of fission, came from the Pu-239 and Pu-241 that was produced by the capture of neutrons in the fertile material, U-238, that was present. (See Table 31.12.)

Most of the reprocessing to date has been performed for the weapons program. Currently, only in England and France are there plants for commercial reprocessing. Japan is building a plant, but for the present is shipping its spent fuel to England and France for reprocessing. Reprocessing provides more efficient use of natural uranium for both the uranium and the thorium cycles. The LMFBR (liquid metal fast breeder reactor) increases the power generated over ten times that discussed above for the once-through cycle.

TABLE 31.12 PWR and BWR Fuel Burn-Up

	Fuel Burn-up (Mwd/MT)	*Fuel Composition (g/MT)*		
		U-235	*Pu-239*	*Pu-241*
	PWR—Plutonium Withdrawal			
Charge	—	32,500	0	0
Discharge	32,000	8,359	5,327	1,213
	PWR—Plutonium Recycle			
Charge	—	6,869	19,849	4,082
Discharge	32,000	2,919	8,765	4,647
	BWR—Plutonium Withdrawal			
Charge	—	25,000	0	0
Discharge	27,000	6,403	4,808	1,034

Dissolution. The spent fuel is dissolved in nitric acid to prepare it for solvent extraction. This is a batch process that is made complex by the fuel cladding and radioactivity. In general the process consists of the following steps:

1. Aging the spent fuel to allow the I-131 to decay.
2. Shearing or chemically dissolving the corrosion-resistant cladding.
3. Completely dissolving the uranium in boiling nitric acid.
4. Treating the off-gas; this includes recycling the acid vapor in a down-draft condenser to trap the tritium, sorbing the iodine on a silver nitrate bed, oxidizing the nitrogen oxides in the residual airborne effluent and scrubbing it with water to remove the nitrogen oxides, and, finally, scrubbing with sodium hydroxide to remove the residual nitrogen oxides.
5. Monitoring the residual airborne effluent and controlling the dissolution to remain below the release specifications.
6. Monitoring meteorological conditions to assure effective dilution and dispersion for release of the residual gaseous effluent to the environment from a stack.

The facilities and equipment are designed for worker and public protection and for accident avoidance. Also, provision is made to allow remote replacement, decontamination, repair, and disposal. The instrumentation used in the process is designed to measure and report all the independent and dependent parameters essential to controlling and monitoring the operation. Management of the defined standard operating procedures, including quality control and quality assurance, is dictated by international standards and safety regulations.

A typical fuel assembly for a PWR contains about 190 half-inch zirconium-clad rods 12 feet long containing in total 1140 lb of UO_2. The assembled rods are mechanically sheared

Fig. 31.18. Purex process showing (top) first solvent extraction cycle, and (bottom) second plutonium solvent extraction cycle.

into 1- to 2-inch lengths and dropped into a perforated basket. This is placed in the dissolver tank where the uranium oxide is dissolved as discussed above. The basket then is removed from the dissolver, and the zirconium hulls and associated hardware therein are dumped into containers, compressed, sealed, and sent to transuranic waste storage. Ultimately, they will go to a disposal facility when one becomes available.

Separation and Purification. In the Purex process discussed here, the uranium, plutonium and fission products are separated by solvent extraction into three different streams (Fig. 31.18). The plutonium stream goes through anion exchange (discussed later) to reduce traces of ruthenium, and the uranium stream goes through silica gel sorption to reduce traces of zirconium. The fission-product stream, which contains the fission products

with the transplutonium products, is sent to waste treatment. About 99 percent of the uranium and plutonium is recovered in separate product streams, and decontamination from fission products by a factor greater than 10^7 is effected.[31]

The spent fuel solution from the dissolver, adjusted to 2 molar nitric acid, flows into the middle of the first column. From there it flows downward countercurrent to a 5 percent tributyl phosphate/95 percent hydrocarbon solvent, which is introduced into the bottom section. The uranium and plutonium transfer to the solvent, leaving most of the fission products in the acid phase, which passes out the bottom of the column. The U/Pu solvent solution is scrubbed in the top half of the column with 3 molar nitric acid to wash out additional fission products.

The valence of the plutonium in the U/Pu solvent stream is reduced from +6 to +4 with hydrazine to lower its extractability, and the stream flows to the middle of the second column, where the plutonium is stripped out of the solvent with mild acid. The uranium passes on in the solvent stream to the third column, where it is washed from the solvent with a weak acid. The fission product stream flowing from the bottom of the first column, containing less than 1 percent of the U and the Pu, is the high-level waste. It is evaporated to remove most of the acid and may be neutralized with sodium hydroxide before going to earth-covered storage tanks.

The uranium and plutonium streams flow separately through second sets of two columns for reextraction and stripping for further separaton from fission products, according to the scheme shown for Pu in the lower portion of Fig. 31.18. The acid waste streams from the second cycles may be recycled to the first to eliminate the second cycle losses.

Pulsed columns are used for these solvent extraction operations in preference to longer packed columns because the latter would complicate the shielding problem. Mixer-settlers also can be used. The solvent is recycled after treatment to remove decomposition products caused by radiation and chemical effects. This treatment may include a combination of distillation with acid and caustic washes.

From the second cycle the plutonium goes through anion exchange for final purification (Fig. 31.19). The principal problem here is due to ruthenium, which is difficult to remove

Fig. 31.19. Plutonium ion exchange flow sheet.

because of its many valence states. The uranium stream goes through silica sorption primarily to remove zirconium, which seems to be carried along as a colloid.

The process is installed in buildings called cayons, where there are rows of concrete shielded cells serviced by overhead gantry cranes for remote replacement and servicing of valves, pumps, piping, and other equipment. Penetrations through the cell walls are offset to prevent radiation from streaming through. Some equipment is located in cells equipped with windows, TV cameras, and manipulators. Cells for remote repair of the equipment may be similarly equipped.

It is occasionally necessary for personnel to enter some of these facilities. This requires considerable time for decontamination and placement of local shielding. The operations that will be performed are carefully planned and rehearsed in order to get the entry over with as swiftly as possible. A special group of employees, the health physicists, are responsible for radiation safety throughout the plant, and they provide close monitoring of these entries.

Another unique consideration is the prevention of nuclear criticality within the cells. In the dissolver and first cycle, criticality is prevented by the presence of the uranium. After the plutonium is separated from the uranium, criticality is prevented by proper design of the vessels and piping. This includes the cell floor and sumps, where materials would collect in case of leakage from the equipment. To prevent criticality the vessels are limited in either diameter or thickness. Vessels and piping are placed in arrangements designed to avoid a critical array.

In one modification of the Purex process, the plutonium is not separated from the uranium. In this version, the first cycle has only two columns instead of three. In addition to reducing the criticality risk, this modification reduces the risk of unauthorized diversion of the plutonium.

RADIOACTIVE WASTE MANAGEMENT

Radioactive waste management involves the treatment, storage, and disposal of liquid, airborne, and solid effluents from the nuclear industry's operations, along with those from other activities that employ the radioactive products. Its strategy involves four approaches: (1) limit generation, (2) delay and decay, (3) concentrate and contain, and (4) dilute and disperse. Combinations of all four of these usually are employed to manage each waste stream.[32]

There are three types of nuclear wastes, based on their radionuclide characteristics:

- Uranium-contaminated waste, principally from mining, milling, and enrichment. Under some circumstances uranium's 15-hour radon decay product requires dditional control. The uranium tails from enrichment generally are not included in waste tabulations because in time they may be incorporated as a fertile material in breeding.
- Plutonium-contaminated waste, principally from weapons materials processing. There are two classes of plutonium wastes: the transuranic, containing mostly plutonium and other transuranics, and the high-level wastes that contain significant heat-generating products, strontium, and cesium.
- Other radionuclide-contaminated waste, generally called low-level waste, as generated in the reactor and fuel cycle, as well as in the radionuclide applications. (See Tables 31.13 and 31.14.)

Limiting the generation of waste is the first and most important consideration in managing radioactive wastes. The Purex process was developed to eliminate the solids additions to the high-level wastes resulting from earlier technology. Improved housekeeping procedures have greatly reduced the low-level and transuranic wastes generated in the power plants, defense plants, and laboratories. Delay and decay is frequently an important strategy because much of theradioactivity in nuclear reactors and accelerators is very short-lived, a few minutes to a few days. Concentrating and containing is the objective of the treatment activities for the

TABLE 31.13 Radionuclides in a PWR Spent Fuel Assembly. Characteristics: 33,000 MWd/t burnup, 14.7 kg U^{235}, 446 kg U^{238}, 108 kg Zr, 10 kg steel and Inconel

			(Curies: years after discharge)			
Nuclide	*Half life (y)*	*mpc (Ci/l)*[a]	*10*	*100*	*1000*	*10,000**
		Zr Clad				
Fe 55	2.7	8E-7	180	10^{-8}	—	—
Co 60	5.3	5E-8	970	0.01	—	—
ZR 93	9E5	6E-8	0.10	0.10	0.10	0.10
		UO_2 Fuel—fission products				
Kr 85	10.4	3E-10	2300	6.8	10^{-25}	—
Rh 106	2.2 hr		270	10^{-25}	—	—
Pd 107	7E6		0.05	0.05	0.05	0.05
I 129	2E7	6E-11	0.01	0.01	0.01	0.01
Xe 133	0.014	3E-10	—	—	—	—
Cs/Ba 137m	30	2E-8	74,000	9300	10^{-5}	
		UO_2 Fuel—transuranium products				
Ra 226	1620	3E-11	10^{-7}	10^{-5}	10^{-3}	10^{-1}
Np 237	2E6	3E-9	7.6	7.6	7.0	3.1
Pu 238	89	5E-8	1000	500	0.5	0.06
Pu 239	24,360	5E-8	142	142	139	108
Pu 240	6580	5E-8	236	235	214	85
Pu 241	13	2E-7	36,000	500	0.01	0.003
Am 241	458	4E-9	770	1700	410	0.005
Cm 244	17.6	7E-9	480	15	10^{-14}	—

[a]Normal human consumption in water would be equivalent to a radiation dose of 500 mrem/year, approximately 2 to 5 times the natural radiation background.
*Note: After 10,000 years the radioactive toxicity is approximately constant for several million years.

longer-lived radioactivity. The concentrate and contain strategy generally involves converting the airborne, liquid, and solid waste to stable solids in corrosion-resistant containers for storage and for transport to and emplacement in carefully selected disposal sites.

Finally, both dilution and dispersion of large volumes of air and water effluents containing very low quantities of radioactivity generally are necessary. The concentration of radioactivity in these effluents is controlled by federal and international regulations, and such effluents are continuously monitored before release to the environment from the waste treatment and other activities. In a particular operation the regulations may be reflected by a set of actions that are triggered by successively higher levels of radioactivity. The first action, where correction by the plant operator is called for, occurs at a small fraction of the permissible level. A somewhat higher level of radioactivity requires both operator action and regulatory notification. The next higher level requires regulatory participation, and if the final level is reached, the operation must be shut down.

Liquid Waste Treatment

Liquid waste usually arrives for treatment as an aqueous solution, and the treatment is primarily concerned with conversion of this solution to a solid form suitable for storage and disposal. Evaporation is the most effective process and achieves decontamination factors for the evaporate of 100 to 100,000 per cycle.

TABLE 31.14 Estimated Cumulative Waste Quantities and Land Requirements—USA[a] (148 GWe in Yr. 2000)

	Volume ($10^6 ft^3$)			*Land (acres)*		
	1980	*1990*	*2000*	*1980*	*1990*	*2000*
Low Level						
Defense	60	80	90	540	650	680
Commercial	15	45	83	150	300	420
TRU[b]						
Defense	0.05	0.05	0.05	20	100	200
Commercial	0.01	0.01	0.01	nil	nil	nil
High-Level[d]						
Defense (sludge & calcine)[c]			0.004	0	0	900
Defense (LLW salt cake)			0.008	nil	nil	15
Commercial spent fuel[d]	0.014	0.7	1.4	20	50	2200
Disposal facilities[d]						
Near surface burial grounds[e]				690	950	1100
Deep geologic vaults				nil	nil	2700

[a]IRG Appendix D, assuming 148 GWe in 2000 and commercial fuel not processed.
[b]TRU in interim surface storage until decision made regarding disposal.
[c]Packaging of defense wastes started in early 1990s.
[d]Spent fuel and defense high-level waste in surface storage until decision made regarding disposal.
[e]Includes scrap from decommissioned facilities.

Fig. 31.20. Slurry-fed vitrification furnace converts aqueous high-level waste into a silica glass form. (*Courtesy USDOE.*)

Ion exchange and scavenging precipitation can be employed at a lower cost than that of evaporation when decontamination factors of 5 to 100 are adequate.

As an example, most of the radioactivity from plutonium production is found in the liquid high-level waste from the first cycle of the Purex process. This liquid is neutralized with sodium hydroxide and stored in earth-shielded tanks. There a sludge settles out that contains most of the radioactivity. The residual liquor is partially evaporated to decrease the volume of the waste, and sodium nitrite crystallizes on top of the sludge. In a process to be used at the Savannah River plant, the sodium liquor fraction, containing most of the cesium fission product, is pumped from the tank and precipitated with tetraphenyl borate. The precipitate will be calcined and packaged for disposal in a high-level repository, and the sodium nitrate crystallized from the residual liquor and sent to a low-level waste repository.

The sludge fraction is slurried from the tank, washed to remove the salt liquor, and then mixed with a pulverized glass fruit. This mixture is to be converted to silica glass in a vitrification furnace, cast in a canister, and sealed. (See Fig. 31.20.)

Airborne Waste Treatment

The largest volume of radioactive waste is made up of the ventilation air from processes, plants, and laboratories. However, the quantity of radioactivity is relatively low. It generally occurs in a particulate form readily removed by dry filters. A special class of paper filters developed for this purpose removes 99.97 percent of 0.3-micron particles. These filters frequently are preceded by scrubbers, sorbers, and roughing filters to extend the life of the paper filter.

A notable exception to the above is the airborne effluent from spent nuclear fuel reprocessing and from tritium production. In this case, iodine-131 volatilizes from fuel reprocessing and must be removed by passing the airborne effluent through a silver catalyst bed, which provides a decontamination factor of about 100,000. Prior to reprocessing, the spent fuel is stored for a sufficient time for decay of the 8-day iodine-131, usually 10 half-lives, to make this treatment adequate. The processing effluent also carries krypton-85, xenon, tritium, and solid radioactive particulates. The particulates are removed by filtration, but the other substances generally can be released. In tritium production, cold traps and sorbers are used to remove tritium from the airborne effluents.

Solid Waste Treatment

The solid wastes from nuclear operations include the concentrates from liquid and airborne treatment along with paper, clothing, lab glassware, and failed equipment. The transuranic and low-level wastes both receive the same but separate treatment. In plutonium operations the first step frequently is monitoring to separate the transuranic from the low-level wastes. Some of these are incinerated to reduce their volume, but this process produces additional airborne and liquid waste. Physical compaction frequently is more useful. Equipment items, including large reactor components, are decontaminated by using a series of acid, chelating, and caustic washes. Finally, the solids may be incorporated in asphalt or concrete to reduce their leachability and then are drummed. Most of these solid wastes go to low-level disposal.

From power reactors with no reprocessing, the high-level waste consists of assemblies of zirconium-clad spent fuel rods to be packaged in stainless steel canisters. If the spent fuel is reprocessed, then the high-level waste will be converted to a silicate glass form similar to that from defense operations. The uranium and plutonium content in the waste, and also possibly the strontium-90 and cesium-137, would be reduced by 90 to 99 percent. The nuclides are useful, and their removal reduces the waste disposal problem several ways; it reduces the toxicity and heat load of the waste and also reduces the incentive for future

generations to undertake the very costly recovery of what may become valuable energy-producing resources.

Storage of Spent Fuel

Spent fuel originally was stored at each nuclear power reactor in water pools designed to accommodate a five-year inventory on the assumption it would be reprocessed later. However, reprocessing did not develop, and there has been a delay in identifying the ultimate repository. The storage pools were reracked to increase their capacity, and now the older reactors also are installing dry storage capacity to accommodate the life of the reactor, approximately 40 years.[33]

In addition, the federal government is committed to begin accepting spent fuel in the late 1990s and is planning to provide away-from-reactor storage that may also include a fuel rod consolidation facility. As is usual for all new nuclear operations, the site and the schedule for this facility will depend on public acceptance.

Waste Disposal

The solidified radioactive wastes for which ultimate disposal must be provided fall into two broad categories: (1) the uranium-containing wastes from milling and the front end of the fuel cycle, including enrichment; (2) the radionuclides and plutonium from the nuclear reactors and spent fuel operations, together with the radionuclides from applications in research, medicine, and other industries.[34]

Agreement has been reached with the public regarding mill waste, and disposal of the accumulated backlog is nearly complete via monitored surface repositories. The proposed technology for disposing of the second category of wastes consists of near-surface burial for low-level wastes and deep mined vaults for the transuranic wastes, including the high-level wastes. The latter approach has encountered strong public opposition, and it is quite possible that the ultimate solution will be surface repositories in somewhat the same manner as for mill wastes.

Mill Wastes. The uranium-containing wastes from milling are mounded and covered with earth. This earth cover prevents erosion and delays for decay the 14-hour radon gas, the gaseous decay product of uranium. These mill waste repositories are located near the mines and mills and are not a very different hazard from the original naturally occurring uranium deposits.[35,36] The depleted uranium from the enrichment operations is stored in cylinders as uranium hexafluoride for future use in the uranium–plutonium breeding cycle. Other uranium wastes from enrichment and fuel fabrication go to the low-level repositories.

Low-Level Wastes. Low-level wastes presently are being entombed in near-surface, earth-covered trenches. These trenches are designed to minimize the inflow of water and drain to a sump that is constantly monitored. Treatment facilities are maintained to process drainage that exceeds licensed levels of radioactivity for release in the groundwater.

The major repository for commercial low-level waste near Barnwell, South California is nearing its capacity. Other repositories around the country were closed when full or when required to be closed by state regulatory action. There have been no health effects or releases of radioactivity exceeding licensed limits from these low-level repositories. The schedule for opening additional low-level repositories has been delayed for several years by the same type of public opposition encountered by proposals relating to treatment and disposal of municipal and other wastes.

Transuranic Waste. It has been proposed that in the future the transuranic (plutonium-containing) wastes be placed in deep mined vaults. This includes the high-level nuclear wastes. The vaults are expected to be about 1000 meters below the earth's surface and will be backfilled and sealed as they are filled. It appears that the repository capacity for high-level wastes will be heat-limited to one kilowatt/acre. This is equivalent to one ton of fuel after storage for ten years. The spent fuel discharged by the year 2000 will require about 2700 subterranean acres. This would

be reduced to 1000 acres if the uranium and the plutonium were removed, and considerably less than that if the strontium and the cesium also were removed. Although reprocessing would reduce the transuranic content by a factor of 10 to 50, this amount would be an insignificant fraction of the transuranic hazard in the waste.

The most restrictive of the current regulations require this waste to be contained for about 10,000 years. Intensive international R&D has been conducted, along with extensive geologic surveys by responsible scientists. A 1990 National Academy of Sciences study concluded that this requirement is beyond the possibility of proof.

Over a billion dollars has been spent on a survey of the continental United States to locate candidate sites for the repository, and several other countries also have conducted site evaluations. Site criteria and regulations have been developed, and several candidate sites have been idntified for more intensive evaluation. In the United States these sites are found in tuff* and salt, and in Germany in salt. Candidate sites in other countries are characterized by granite and clay.

Extensive experimental work has been carried out, and perhaps the most interesting study is related to a group of natural nuclear reactors found by the French in Gabon, Africa. During Neolithic time, while uranium was being deposited in a swamp, the developing ore body repeatedly achieved criticality. This occurred over a period of centuries at a series of interconnected sites. It appears that the long-lived products were closely held within the ore body. Unfortunately, the translation of this information to other geologic environments poses many uncertainties.

The research on repositories was started in the late 1950s, and in the late 1960s a pilot study using spent fuel was successfully completed in a Kansas salt mine. However, by that time public opinion forced the program back into increasingly extensive and costly exploration and experimentation activities. The nuclear utilities are now being assessed 0.1 mil per kW electric that is expected to pay for disposal of their high-level spent fuel wastes.

*Rock composed of volcanic detritus usually fused together by heat.

TABLE 31.15 Conceptual Methods for Radioactive Waste Disposal

Terrestrial	Shallow land burial
	Deep geological vaults—continental sites
	—island sites
	Caves
	Deep hole
	Sub sea bed
	Ice cap
	Hydrofracturing
	In situ rock melting
Extraterrestrial	Solar orbit —Space shuttle
	Moon crater—Rocket/soft lander
	Solar escape—Electric cannon
Transmutation	Fission reactors
	Fusion reactors
	Electromagnetic accelerators

In 1989 a deep mined repository, The Waste Isolation Pilot Plant, was opened in a bedded salt formation near Carlsbad, New Mexico. However, state objections are delaying the experimental emplacement of transuranic wastes from the defense program. In addition, a Nevada site with a tuff formation above the water table has been selected to be intensively studied for disposal of high-level transuranic wastes. In this case also, state objections are delaying the work. It is very uncertain when the first repository will be receiving any transuranic waste, and it probably will not before 2010.[37]

Over these years of study, many alternatives have been considered (Table 31.15). Of these, the subsea bed appears to warrant further attention, but the reliability of sea transport to the site could be limiting. There are several large areas of the ocean bottom covered with desertlike deposits of silt several kilometers deep.

TRANSPORTATION OF NUCLEAR MATERIALS

The transport of nuclear materials reflects many years of experience based on millions

of shipments. These shipments range from microcuries in liquid and gaseous forms in glass vials carefully packaged in sealed cans inside cardboard boxes to megacurie quantities contained in spent fuel packaged in hundred-ton shipping casks. They are shipped by both private and public carriers, by car, bus, truck, railroad, airplane, barge, and ship.

During these years there have been many incidents involving all forms of transportation, both commercial and military, and a few of these events resulted in some contamination of facilities. However, there have been no health effects to any individual transport worker or the general public.

Shipments of radioactive materials, originated by many different shippers throughout the world, are controlled by an established code of national and international regulations. Although most nuclear operations relate to a localized population, transport potentially affects large numbers of people. Many communities have attempted unsuccessfully to prohibit nuclear transport through their jurisdiction. Some, notably New York City, have succeeded.

Shipments of spent nuclear fuel constitute the largest quantity of nuclear materials being transported. These shipments involve 50-ton truck casks with the capacity for one to several fuel assemblies. Rail and ship casks weighing about 100 tons carry many assemblies. The design of these casks provides:

1. Radiation shielding to limit radiation to the workers and the public.
2. Passive cooling for the heat-generating spent fuel, up to several kW per assembly.
3. Assured containment in the event of the maximum accident.

Current casks are designed to handle five-year aged spent fuel that travels by truck, rail, and ship. There are frequent ocean shipments from Japan to England and considerable rail shipment in Europe. Most of the spent fuel shipments in the United States are by truck. No cask failure has been experienced in any accident to date.

The transport of plutonium wastes to the transuranic and high-level repositories will represent a major increase in the quantities of radioactive materials shipped; so an extensive demographic study has been made to examine potential exposures of the public and to select routes to minimize such exposure.[38,39]

New casks are being developed to maximize the efficiency of the longer-aged spent fuel shipments and to reduce the exposure of the workers involved in loading and unloading the casks. There are expected to be two families of casks, those transporting the spent fuel from the reactors to a temporary storage facility, primarily truck casks, and those transporting the spent fuel from the storage facility to the repository, primarily rail casks. Another set of casks is being developed for the high-level defense waste at the Hanford and Savannah River facilities.

THE NUCLEAR REACTOR

Development of the commercial nuclear power industry for the generation of electricity began in the early 1950s. The first commercial electric power was produced using a light water reactor (LWR) at Shippingport, PA in 1957. By the 1970s large numbers of power reactors had come on-line. In the United States all but one were LWRs, but in Canada the preference was for heavy water reactors (HWRs). There were many gas-cooled graphite reactors (GCRs) in England, France, and Russia and one in the United States. France and Russia had the first liquid metal fast breeder reactors (LMFBRs) operating, with several engineering test reactors in the United States. The characteristics of typical units are presented in Table 31.16. Nuclear reactors can be classified in several ways:

- *Burners vs. breeders:* Burners depend on uranium from ore but may recycle the residual fissionable material from the spent fuel to reduce the ore requirement. Breeders produce more fissionable material than they consume by converting either uranium-238 into plutonium-239

TABLE 31.16 Characteristics of Nuclear Power Reactors

	PWR[a]	*BWR*[b]	*Candu*[c]	*LMFBR*[d]	*HTGR*[e]	*AGR*[f]
Electric power (GWe)	1.1	1.06	0.74	0.23	0.33	0.63
Plant efficiency (%)	33	33	30	40	42	42
Fuel	UO_2	UO_2	UO_2	UO_2	UO_2ThO_2	UO_2
Clad	Zr	Zr	Zr	SS	C	Mg
Uranium loading (t)	99	155	130		95	
U-235(5)	3.0	2.9	0.81	12*	40	0.75
Burnup (GWd/t)	30	28	7.3	100	110	
SWU(t/GWey)	109	114	0	0,03	130	
Fueling internal (y)	1.1	1	0.003	0.5	1	
Fueling fraction	0.3	0.2	0.003	0.3	0.3	
Power density (kW/kgU)	38	25	26	160	115	38
Coolant/Moderator	H_2O	H_2O	D_2O	Sodium	He/C	CO_2
Pressure (bars)	155	73	100	2	45	40
Temp (°C)	327	—	310	560	750	670
Flowrate (t/s)	21	15	12	3	0.4	
Conversion ratio	0.6	0.6	0.6	1.2	0.8	0.6

*Including Pu-239
[a]Zion 1 1973.
[b]Peach Bottom 2 1974.
[c]Bruce 1 1977.
[d]Phenix 1973.
[e]Fort St. Varain 1979.
[f]Hinkley Point B1 1976.

or thorium-232 into uranium-233. The breeders depend on the plutonium-239 or uranium-233 from the burners to bring them into equilibrium, and this may take 30 years.

- *Choice of neutron moderator—light water (protium), heavy water (deuterium), sodium (liquid metal), or carbon (graphite):* Except in the case of graphite, the moderator also serves as the coolant.
- *Neutron energy (thermal vs. fast):* The sodium moderated reactor operates with fast neutrons to breed using the uranium cycle.
- *Pressurized vs. boiling LWRs:* The pressurized water reactor (PWR) transfers its energy from the fuel to an intermediate heat exchanger to generate the steam that drives the turbine. The boiling reactor generates steam within the reactor that goes directly to the turbine.

The PWR, CANDU, and graphite reactors are the major producers of electricity from nuclear energy. When compared on the basis of cost, reliability and safety, they are competitive. The PWR is the most widely used, but there are several BWRs in use. In 1990 there were 16 CANDUs operating in Canada where they were developed, with several in other countries as well.

TABLE 31.17 PWR and CANDU Requirements for Uranium from Ore

	Metric Tons/GWy-electric	
	PWR	*CANDU*
Once-through cycle	217(1)	175(2)
Pu/U recycle	157	129(3)

(1)3.25 percent U-235 fuel.
(2)0.71 percent U-235 fuel (natural uranium).
(3)Pu/U recycled to CANDUs, from PWR once-through spent fuel (PWR/TANDEM).

When PWRs and CANDUs are compared on the basis of uranium utilization, the CANDU requires 19 percent less uranium from ore (see Table 31.17). As there is now an abundance of low-cost uranium from ore and enrichment service, the electric utilities continue to favor the PWR and BWR types. Operating experience with graphite reactors in England, France, and Russia also has been generally satisfactory. However, the graphite moderator represents a large inventory of

combustible material, which contributed significantly to the severity of the Chernobyl accident. Further use of this technology is unlikely.

A new family of nuclear reactors has been designed by the PWR, BWR and CANDU suppliers, which are now (1992) being licensed. It is the industry's objective to use proven technology to provide reactors that are simpler and less costly to build than older designs. These reactors would have wider safety margins through the increased use of passive safety systems, relying more on reactor design and less on operator action to prevent accidents. They would be made less complex by drastically reducing the number of components necessary for operation. This design also would facilitate plant construction by requiring less concrete, fewer valves and pumps, and less piping. Most likely the future will see introduction of modular built systems that will be prefabricated at the factory. Safety considerations always will be paramount.[40,41]

Light Water Reactors

Pressurized Water Reactor. The PWR contains three coolant systems: (1) the primary system, which removes heat from the reactor and partially controls nuclear criticality; (2) the secondary system, which transfers the heat from the primary system via the steam generator to the turbine-driven electric generator; (3) the service water system (the heat sink), which dumps the residual coolant energy from the turbine condenser to the environment. The service water is recirculated from a river, lake, ocean, or cooling tower. In the primary system (Fig. 31.21), dissolved boron is present to control nuclear criticality. Fixed-bed ion exchange units are used to maintain the water quality in both the primary and the secondary systems. In addition, the chemical and volume control system reduces boron concentration during the power cycle to compensate for fuel burnup. These operations are carried out continuously through bypass systems.[42] A more complete

Fig. 31.21. Pressurized water reactor (PWR) for power generation. (*Courtesy USDOE.*)

Fig. 31.22. Boiling water reactor (BWR) for power generation. (*Courtesy USDOE.*)

view of the current technology is presented in Table 31.16.

Boiling Water Reactor. In the BWR, steam is generated in the reactor and goes directly to the turbine (Fig. 31.22). Although the BWR eliminates the intermediate heat exchanger, it places the steam turbine in a radioactive environment. Here in the primary system, there are full-flow ion exchange units treating about 30,000 gal of water/min, which serves two objectives: (1) removal of radioactivity from neutron-activated corrosion products and fuel cladding leaks, and (2) corrosion protection of the reactor, fuel, and turbine from possible leakage of service water to the primary coolant from the turbine condenser.

In the following paragraphs, three important aspects of LWRs are discussed: fuel elements, the primary coolant system, and reactor containment. They each play an important role in providing multiple barriers that prevent radioactive fuel and fission products from being released to the environment.

Fuel Elements. The fuel element is designed to provide the primary containment for the radioactive fuel and fission products over the three- to four-and-one-half-year operating life of the PWR and BWR fuel. In addition, without significant problems, it has provided the containment of the spent fuel after discharge from the reactors for more than ten years under water in the storage basins at the reactors. At some of the older reactors it is now being packaged in stainless steel canisters and placed in dry storage vaults.[43]

The quality of the reactor coolant depends on the integrity of about 80 miles of 30-mil Zircaloy-4 tubing operating at temperatures up to 350 °C. The PWR fuel assembly, which is about 16 feet long, is made up of about 190 fuel pins. Each fuel pin contains fuel in the form of ceramic uranium oxide pellets in a Zircaloy-4 tube with a 24-mil wall. This tube is the first barrier against releasing radioactive fission products to the environment. (See Fig. 31.23.)

A typical PWR contains about 110 tons of fuel charged with 3.4 percent U-235. The economics of the fuel cycle requires this degree of enrichment to achieve a reasonably long operating life and high energy yield from each assembly. Because these power stations

represent a significant part of the power system capacity, frequent shutdowns are undesirable. With the increased experience and expectation of a once-through fuel cycle, the enrichment is being increased to raise the energy yield to 50,000 MWdays/metric ton and to extend the time between refueling shutdowns beyond 18 months. This reduces the amount of replacement power needed during shutdowns, which generally comes from the older, more costly fossil-fuel-powered plants or must be purchased from outside the system. Approximately one-third of the fuel is replaced in the core of the reactor during each refueling shutdown. The new fuel is placed in the outer regions of the core to flatten the power distribution in the reactor core, and at the same time greater burnup is achieved by placing the partially spent fuel toward the center of the core where the neutron flux is higher.

A 1000 MWe BWR core contains about 165 tons of UO_2, which is charged with 2.5 percent U-235 (Fig. 31.24). Here the integrity of the 32-mil fuel cladding (Fig. 31.24) is even more important than in the PWR because the steam is generated in the reactor core and goes directly to the turbine. This has proved to be a feasible design and eliminates the PWR expense of replacing steam generators.

Fig. 31.23. Cutaway of PWR fuel element with the control rod cluster assembly. Element contains 1140 lb UO_2 in 193 rods. (*Courtesy USDOE.*)

Fig. 31.24. Cutaway of BWR fuel element containing 488 lb UO_2 in 49 rods. (*Courtesy USDOE.*)

Two engineering system demonstrations were performed to reduce the uranium-from-ore requirements of LWRs: (1) recycle of the plutonium and (2) conversion to the thorium–uranium cycle to achieve thermal breeding. The demonstration phase of the plutonium recycle development was carried out in seven power reactors. Several LWRs originally were started up on the thorium–uranium cycle, and a light water breeder reactor (LWBR) core using the thorium cycle was tested in the Shippingport reactor. Although these tests did not resolve all the technology questions, no major problems were uncovered. In the United States, plans for fuel recycling and the use of the thorium cycle have been postponed because of the low cost of uranium from ore as well as the perceived danger of diversion of the fissionable materials.

Primary Coolant System. Nuclear technology faces challenges on most of the frontiers of engineering, with design and fabrication of the primary coolant system being one of the greatest. The reactor vessel and associated piping and heat exchangers comprise the second barrier to prevent the release of radioactive materials from the fuel

to the environment. The steel for the reactor vessels must meet metallurgical standards developed for them to withstand the temperature and pressure cycles as well as neutron irradiation, which in time reduces the ductility of the metal. Carbides and phosphides precipitated in the grain boundaries prevent slippage along the crystal planes. The design criteria for optimum performance require operating pressures of 1050 psia for BWRs and 2250 psia for PWRs. Typical dimensions for vessels serving 1000 MWe stations range from 21 feet in diameter by 70 feet in height for BWRs, to 14 by 42 feet for PWRs, with vessel weights of 782 and 459 tons, respectively. These vessels are shop-fabricated, and their transport to the reactor site tests the ingenuity of the carriers.[44,45]

Containment. Under all conditions of normal operation and maximum credible accidents, any release of radioactivity from the nuclear reactor to the environment must be below acceptable limits. The maximum credible accident in the LWR is postulated to be the loss-of-coolant accident (LOCA). This could, in the absence of the redundant engineered safety systems, result in meltdown of the fuel with release of the radioactive fission products. The fission products in turn could be transported out of the core in the escaping vaporized coolant. To prevent escape of the radioactive coolant from the facility, a containment vessel is built around the primary coolant system. In addition, engineered safety systems rapidly shut down the nuclear reaction, flood the reactor vessel, condense the vapor, and scrub out and concentrate the radioactivity. One PWR approach to accomplishing this uses a large-surface ice condenser. The ice is always present during reactor operation. In the BWR (Fig. 31.22), the escaping steam is vented into a vapor suppression chamber where it is condensed. In both reactor systems, there are redundant systems for cooling the reactor and reactor containment.

As anticipated, there have been occasional equipment failures involving reactors, but the safety systems have been sufficiently redundant so that one or more always have worked. Even in the Three Mile Island accident in 1979, the safety systems worked. Much of the damage resulted from operator actions to override the safety systems. As concluded in the Reactor Safety Report, the limitations of the operator created and then seriously aggravated the Three Mile Island incident. Nevertheless, the features of the containment system prevented significant exposure to the plant personnel or any off-site individual, this despite failure of the barriers provided by the fuel cladding and the primary coolant system.

LWR tests-to-failure had been performed to evaluate accident scenarios involving LOCA events such as occurred in the Three Mile Island incident. The power burst tests in a 20 MWt PWR have created fuel failures and defined the initiating conditions. The LOCA tests with a 50 MWt PWR have demonstrated recovery from catastrophic major feedwater and steam line breaks without fuel damage.

CANDU Heavy Water Reactor

This Canadian reactor design utilizes (1) natural uranium as the fuel, thus avoiding the necessity for uranium enrichment, and (2) heavy water as both coolant and moderator, to maximize neutron utilization. In addition, this reactor can efficiently produce cobalt-60, which has commercial markets for product irradiation. The CANDU requires 18 percent less uranium from ore than the LWR once-through cycles need. The spent fuel contains 0.23 percent U-235, roughly equivalent to the tails from the gaseous diffusion enrichment operation. It also contains 0.27 percent fissionable plutonium. The average residence time of a fuel is about one year, with approximately 0.3 percent replaced each day. This requires refueling during operation. The CANDU fuel bundles contain 37 Zircaloy-clad UO slugs, a less complex design than that of the LWR assemblies. Many alternative fuel cycles are being considered for the CANDU to further decrease the requirements of uranium from ore, but they will require enrichment or reprocessing.[46]

Liquid Metal Fast Breeder Reactor

The salient features of the liquid metal fast breeder reactor (LMFBR) include a fuel doubling time of 10 to 30 years, along with a high coolant temperature for more efficient energy conversion and low operating pressure through the use of liquid sodium as the coolant. Operation of the reactor with fast neutrons increases the neutron multiplication factor for the mixed uranium–plutonium oxide fuel. The power density in the LMFBR is about five times greater than that in the LWR cores; so the reactor vessel is much smaller. This core is surrounded with radial and axial blankets of fertile material. These blanket regions, which capture the leakage neutrons, must be used to achieve breeding, which consists of the production of fissionable material from fertile material. Finally, the use of the uranium–plutonium fuel cycle makes optimum use of the plutonium produced in the LWRs.

All major countries are participating in LMFBR demonstration plants, ranging from about 200 to 500 MWe. In the United States, the Experimental Breeder Reactor I (EBR I), the starting point of this program, went critical in 1951. LMFBRs are operating in the United States, England, France, Japan, and Russia. In France, the Phenix, a 590 MWt/230 MWe reactor, has operated since 1973, and the Super Phenix, a 1200 MWe reactor, first generated power in 1987 and is now in commercial operation. In the former Soviet Union, the BN-350 MWe reactor has been producing power and steam heat for Shevchenko on the shore of the Caspian Sea since 1973; and another reactor, the BN-600, has been operating since 1980. In the United States, the 400 MWt Fast Flux Test Facility started operation in 1980 to test fuels and materials. The objectives of this program were satisfied in 1989, and it may be shut down.

During the 1980s the Argonne National Laboratory was developing a metal-fueled LMFBR in contrast to the oxide-fueled version described above. The principal advantage of the metal fuel is the ease of recycle using electrometallurgical technology. The program is called the Integral Fast Reactor with emphasis on high burnup instead of breeding, possibly 185,000 MW-thermal/metric ton versus 33,000 for the present-day PWRs (Fig. 31.25.) Early work with metal fuel showed that it expanded because of the presence of the xenon and ruptured the cladding. The fuel rods have been redesigned to accommodate this phenomenon. Finally, each reactor site would have its own reprocessing unit, and this should greatly reduce concern about proliferation. Cost estimates indicate that the cost of the electricity generated might be equivalent to that from LWR and HWR systems. Much work remains to be done to confirm all these attractive indications.[47]

Other Nuclear Reactors

Many other nuclear reactors have been developed for research, engineering development, nuclide production, and mobile power. Must noteworthy are the nuclear propulsion systems for naval applications. The first nuclear submarine, the *Nautilus*, was commissioned in 1954. Today there are more than one hundred nuclear submarines, along with a number of aircraft carriers and other surface vessels. In the United States, the PWR technology for electric power generation is an outgrowth of the nuclear naval development program. Other nuclear propulsion systems have been studied for arctic tractor trains, aircraft, and rockets for space propulsion, but these studies were terminated before completion owing to insufficient benefit. Recently, nuclear reactor development was started again to provide power on orbiting space stations in connection with the United States Strategic Defense Initiative program. As part of this activity, a "space" reactor was purchased from the Russians who have been employing them for many years.

The remaining classes of nuclear reactors range from zero-power, subcritical neutron sources for university training to large-scale reactor systems for plutonium-239 production. Portable reactors have provided heat, power, and water to U.S. bases in Alaska, Antarctica,

Fig. 31.25. Schematic of integral fast reactor and fuel cycle concept (IFR). (*Courtesy USDOE.*)

and Panama. Private industry has operated various test reactors for reactor studied and radio-isotope production.

RADIATION PROCESSING

Next to the generation of electric power, radiation processing is potentially the most important commercial application of nuclear energy. Radiation processes have been developed for treating food and medical supplies to inhibit growth of bacteria, viruses, fungi, and insects, and for polymerization of plastics and rubber. They take the place of thermal and chemical treatment to achieve these objectives. Radiation also is an important tool in the medical treatment of cancer. However, what may be one of its most important applications, extending the storage life of foods, has been seriously delayed by public concern, in spite of extensive tests showing that irradiated foods meet all international safety standards.

In early radiation processing work, cobalt-60 was the principal source of radiation, but now it has been displaced by machine-produced electrons wherever possible. Cobalt-60 is commercially available from the Canadian CANDU power reactors; each reactor can produce 30 to 50 megacuries annually, which at a dollar per curie more than compensates for the neutron cost.

Commercial information regarding the application of machine-produced electrons is closely held, but the following are a few examples of their use:

1. Low energy electrons, up to 0.5 MeV, are in general use to polymerize thin plastic sheeting and coatings. Fixed beams up to 2.5 meters wide irradiate 0.4-mm-thick

products with 10 kiloGray (kGy) at 1600 meters/hour.

2. Electrons with energies up to 5 MeV are widely used in scanned beam machines. The typical scan width is 1 to 2 meters when applied to the polymerization of thicker materials. These machines also are being used for irradiating grain and chicken feed. The Russians are treating 200 tons/hour of grain with a dose of 0.4 kGy, and the Israelis treat chicken feed with 3.8 kGy at 15 tons/hour. In both cases the flowing-bed depth is about 6 mm.
3. A 5 kW beam of 7 MeV electrons is delivering a 3 kGy dose to flat cakes of deep frozen cakes of chicken meat at a rate of 50 tons/day; disposable medical supplies are being treated at the rate of 100 cubic meters/day.[53]

RADIOISOTOPE APPLICATIONS

The tremendous release of energy from nuclear reactions makes possible a unique family of applications for long-lived radioisotopes that are important to health, science, and industry. Whereas fission and fusion occur almost instantaneously, other radioactive decay processes occur in times ranging from a few minutes to thousands of years. The general areas of application may be grouped into irradiation, thermal energy generation, and tracer applications.[50]

Radiation Sources

Radiation from radioactive nuclides is used to detect changes in density or other characteristics of materials, to promote chemical, physical, or biological changes, and to provide a source of thermal energy. Some radioactive materials find use in industry, research, and medicine as tracers for physical, chemical, and biological processes.[52]

The very important radioisotope cobalt-60 has a 5-year half-life and emits 1.3 MeV gamma rays. (Co-60 is made by irradiating naturally occurring Co-59 with neutrons in a reactor.) It is the major radioisotope in a number of applications including the following:

1. Destruction of malignant tumors.
2. Sterilization of prepackaged medical supplies, particularly those that are heat-sensitive or can be contaminated by chemical agents.
3. Food irradiation to extend shelf life.
4. Radiographing metals to determine weld quality. (Iridium-192, with a 0.3 MeV gamma ray, provides better definition than Co-60 in some cases but has a shorter half-life.)
5. Polymerization of coatings on electrical conductors and paper and monomers impregnated into wood and concrete. (Machine radiation generally is preferred in applications involving thin materials.)

There also are many applications for alpha and beta radiation sources, including:

1. Polonium-210 alpha activation of beryllium to yield neutrons that in turn are used to start up nuclear reactors.
2. Plutonium-238 activation of beryllium to produce neutrons for logging drill holes to detect hydrogenous materials in geologic formations.
3. Strontium-90 beta sources to measure the thickness of paper in paper manufacture and also to discharge static electricity by ionizing the air in high-speed printing operations.

Radioisotope Thermoelectric Generators

These devices (RTGs) provide reliable long-life sources of electrical energy. The thermal energy required to drive the thermoelectric element is provided by a long-lived radionuclide. The electricity is generated by passing the thermal energy through the thermoelectric element from the heat source at one end (the hot foot) to a heat sink at the other end (the cold food). Germanium silicide is now the thermoelectric material of choice; because of its higher thermal stability, it can be driven

Fig. 31.26. The multi-hundred-watt electric generator. Three units on *Voyager* generated 475 watts from the decay heat of plutonium-238 acting on thermoelectric couples. (*Courtesy USDOE.*)

with larger temperature differentials. Lead telluride is more efficient per degree of temperature difference but has lower thermal stability.[54]

Plutonium-238 is the most important heat source; it is an alpha emitter whose radioactive decay energy is easily sorbed and converted into thermal energy. It also is easily shielded.

In addition, its long half-life, 90 years, makes possible hundred-watt power sources with a design life exceeding ten years. Although solar cells generally comprise the technology of choice to power extraterrestrial missions, a Pu-238 RTG was used for the Apollo moon landings. Such devices are used when the mission path is shielded from the sun, and they continue to be used for deep-space missions. (See Fig. 31.26.)

Terrestrial RTGs have also been developed and are commercially available. A beta emitter, 30-year strontium-90, has been the major radionuclide used for terrestrial applications. Although beta particles (negative electrons) are easily sorbed and converted to thermal energy, the sorption process is accompanied by the emission of X rays that require shielding. This necessitates a heavy device, about 300 lb for a 5-watt source. Because strontium-90 is a fission product, the supply now is limited to 100 megacuries recovered in the 1960s from the Hanford defense waste. This is equivalent to about 10 kW of electrical energy, assuming a 5 percent conversion efficiency.

Tracers

The radiation emitted from decay of single atoms can be easily measured. Thus, small amounts of radioisotopes can be incorporated into systems to study, for example, chemical reactions and the flow of fluids through complex systems such as organs of the human body. They also can be used in more mundane applications such as locating leaks in buried piping. Many organic compounds tagged with carbon-14 in specific positions are commercially available.

Because C-14 with a half-life of 5700 years occurs in nature, it can be used to determine the age of carbon-containing artifacts. The basis for such determinations is the loss of C-14 that was trapped in the artifact (animal, human, fired pottery, and so on). The related art and science of carbon dating are fairly complex and thus subject to significant uncertainties.

TABLE 31.18 Radionuclides Used in Nuclear Medicine

For organ function and tumor visualization:	
Technetium-99m	Carbon-11
Iodine-131	Oxygen-15
Gallium-67	Nitrogen-13
Strontium-85	Fluorine-18
Indium-111	Iodine-123
Xenon-133	
For radiation of tumors:	
Cobalt-60	Cesium-137
Radium-226	Iodine-131

Nuclear Medicine

Nuclear medicine is one part of the nuclear readily accepted by the public. However, there is strong opposition to the siting of the low-level waste repositories needed for disposal of the wastes that result from the practice of nuclear medicine.

The medical applications of nuclear technology range from in vitro and in vivo injections for diagnostic tests to cobalt radiation for cancer therapy. A new medical speciality was created, a family of compact cyclotrons was developed to provide short-lived nuclides, and a sizable industry evolved to produce technetium. Until the nuclear industry was created, technetium had been missing from the chart of chemical elements because the half-life of the most stable member was too short, 21,000 years. Technetium and several other nuclides of importance here are discussed elsewhere in the chapter in connection with their production. (See Table 31.18.)[49,51]

The two principal applications of nuclear medicine are for radiation treatment of tumors and for visualizing organ functions and tumors. These nuclear-medical procedures supplement other medical treatment and diagnostic procedures, and are available in most hospitals throughout the industrialized countries and in the major hospitals in other countries.

REFERENCES

STATUS AND OUTLOOK

1. Finger, H., "Need for Nuclear Energy, National Energy Strategy Hearings," US Council for Energy Awareness, Washington, DC, 1989.
2. "Nuclear Electricity for September 1990," *Nucleonics Week*, **45**, NY, Nov. 8, 1990.
3. Reactor Safety Study, WASH 1400 (NUREG 75/014) US Regulatory Regulatory Commission, Washington, DC, 1975.
4. International Fuel Cycle Evaluation, INFCE/PC/2/9. International Atomic Energy Agency, Vienna, 1980.
5. Code of Federal Regulations, Title 10 Energy, Chapter 1 Nuclear Regulatory Commission, Washington DC.
6. "The Accident at Three Mile Island," Staff Report to the President's Commission, Nuclear Regulatory Commission, Washington DC, 1979.
7. "Chernobyl Nuclear Power Plant Accident—Health and Environmental Consequences," DOE/ER-0332 US Department of Energy, June 1987.
8. Uranium Resources, Production and Demand, OECD Nuclear Energy Agency and International Atomic Energy Agency, Paris, 1990.
9. Annual Report to Congress 1988, Energy Information Administration, Washington DC, 20585, 1989.
10. Annual Energy Review 1988, DOE/EIA-0384[88], Energy Information Adminstration, Washington DC, 20585.
11. "Energy Efficiency," *The Energy Daily*, 18, **64**, Washington DC, 1990.
12. Parent, J. D., A Survey of United States and Total World Production, Reserves and Remaining Recoverable Fossil Fuel and Uranium, Institute of Gas Technology, Chicago, IL, 1977.
13. Friedlander, G., Kennedy, J., Macias, E., Miller, J., *Nuclear and Radiochemistry*, 3rd Ed, J. Wiley & Sons, New York, 1981.
14. *Nuclides and Isotopes*, 14th Ed, 202-637-4000 GE Nuclear Operations, 175 Curtner Ave., M/C397, San Jose, CA, 95125.

Radioisotope Decay

15. A Status Report on Controlled Thermonuclear Fusion, STU/PUB/872 International Fusion Research Council, International Atomic Energy Agency, Vienna, 1990.
16. Starpower, the U.S. and the International Quest for Fusion Energy, OTA-E-338, Office of Technology Assessment, Congress of the United States, Oct. 1987.
17. Kauffman, G., "Beyond Uranium, Chemical and Engineering News," **47**, Washington, DC, Nov. 1990.
18. Crandall, J., "The Savannah River High Flux Demonstration, USAEC Report DP999," US Atomic Energy Commission, Washington, DC.
19. Lagunas-Solar, M., "Cyclotron Production of No-Carrier-Added Medical Radionuclides," 7th Conference on the Applications of Accelerators in Research and Industry, Denton TX, 1982.
20. Meyer, W., and Plascjak, P., Cyclotron Group, National Institute of Health—personal communication.
21. Stable Isotopes for Research and Industry, ISOTEC, Maimisburg OH 45342, 1990.
22. Separation and Applications of Stable Isotopes, Avona and Spicer, American Laboratory, April 1987.
23. "Separation of Hydrogen Isotopes," Rae, H., Editor, American Chemical Society, Washington, DC, 1978.
24. Fuel Cycle Review 1990, Nuclear Engineering International.
25. Eister, W., Stoughton, R., Sullivan, W., *Processing of Nuclear Reactor Fuel, Principles of Nuclear Reactor Engineering*, Glasstone, S., (Ed.), Van Nostrand, NY, 1955.
26. Leist, N., "The Benefication and Refining of Uranium Concentrates," NLCO 1067, US Atomic Energy Commission, 1970.

Fuel Production

27. Mantz, E., "Production of Uranium Tetrafluoride and Uranium Metal," USAEC Report NCLO 1068, U.S. Atomic Energy Commission, 1970.
28. U.S. Uranium Enrichment, The Case for Restructuring, The council for Energy Awareness, Washington, DC, 1988.
29. Olander, D., "The Gas Centrifuge," *Scientific American*, V. 239, No. 2, 1978.
30. Long, J., *Engineering for Nuclear Fuel Reprocessing*, Gordon and Breach, New York.
31. *Solvent Extraction and Ion Exchange*, Logsdail, D., et al., J. Wiley & Sons, New York, 1985.
32. "Integrated Data base for 1989: Spent Fuel and Radioactive Waste Inventories," Department of Energy, Washington DC, 20585, 1989.
33. "Dry Storage Casks for Spent Nuclear Fuel," *Journal of the Institute of Nuclear Materials Management*, Northbrook, IL, 60062, May 1990.

34. Annual Report to Congress, Office of Civilian Radioactive Waste Management, U.S. Department of Energy, washington, DC, 1989.
35. Environmental Assessment of Remedial Action at the Monument Valley Uranium Mill Tailings Site, Monument Valley, AZ, U.S. Department of Energy, Washington, DC, 1989.
36. Uranium Mill Tailings Remedial Action Program, Annual Report, 1989, Washington, DC, 1989.
37. Eister, W., "Materials Considerations in Radioactive Waste Storage," p. 6, *Nuclear Technology*, **1**, Jan. 1977.
38. Rail Transportation Corridor Analysis, BMI/ONWI-617, U.S. Department of Energy, 1986.
39. Transporting Spent Nuclear Fuel, U.S. Department of Energy, Office of Civilian Radioactive Waste Management, Washington, DC, 1986.
40. Benedict, M., Pigford, T., Levi, *Nuclear Chemical Engineering*, McGraw-Hill, New York.
41. "Nuclear Reactors and Technology," Monthly Abstracts, Office of Scientific and Technical Information, Department of Energy, Oak Ridge, TN.
42. "International Comparison of LWR Performance," Beckjord, E., et al., Report MIT EL 87-004, Masschusetts Institute of Technology, Cambridge, MA, Feb. 1989.
43. Frost, B., *Nuclear Fuel Elements*, Pergamon Press, New York, 1982.
44. Cohen *Water Coolant Technology of Power Reactors*, Gordon and Breach, 1969.
45. Advanced LWRs, Nuclear Industries, July 1988, U.S. Council for Energy Awareness, Washington, DC.
46. *A Catalogue of Advanced Fuel Cycles in CANDU-PHW Reactors*, Veeder and Didsbury, Chalk River Nuclear Laboratories, Chalk River, Ontario, Canada, KOJ 1JO, 1985.
47. Till, C., Chang, Y., "The Integral Fast Reactor," *Advances in Nuclear Science*, **20**, 1988.
48. Glasstone, S., The Effects of Nuclear Weapons, U.S. Government Printing Office, 1977.
49. Eister, W., et al., Radiopharmaceuticals and Short-Lived Radioisotopes, Radioisotope Production Study, ASao Paulo, Brazil, IAEA-124, International Atomic Energy Agency, Vienna, 1970.
50. Eister, W., et al., Radioisotope Production in the U.S., (see above).
51. Eister, W., et al., Radioisotope Generators America, (see above).
52. Eichholz, G., *Radioisotope Engineering*, Marcel Dekker, New York, 1972.
53. "Machine Sources for Food Irradiation," U.S. Department of Energy, Washington, DC, 1988.
54. *Handbook of Isotopic Power Source Characteristics*, Arnold, ORNL 3576, Oak Ridge National Laboratory, Oak Ridge TN, 1964.

Index

Abietic acid, 265–268
Acetaldehyde, 821
Acetal resins, 650
Acetic acid
  by fermentation, 937
  from acetaldehyde, 823
  from butane, 840
  from methanol, 804, 806
  from wood, 260, 263, 270
Acetic anhydride, 245, 250, 823
  from coal, 590
Acetone
  by cumene oxidation, 834
  from molasses, 952
  from wood, 260, 263, 270
Acetylene, 456, 808
Acetylsalicyclic acid, 1002
Acid black dyes, 896–899
Acid Blue, 25, 966
Acid green dyes, 904, 906
Acid Orange, 7, 894
Acid red dyes, 894–897, 904, 905
Acid Yellow, 23, 895
Acid yellow dyes, 867, 895, 904
Acridines, 247
Acridine yellow, 905
Acrolein, 835
Acrylic acid, 729, 835
Acrylic adhesives, 727
  acrylate synthesis, 728
  methacrylate synthesis, 729
Acrylic esters, 835
Acrylic fibers, 737, 771
Acrylic resins, 642
Acrylonitrile, 250, 616, 831
Activated sludge process, 35
Acylglycerols, 275
Adhesion, 708
Adhesives, 708
  acrylic, 727
  epoxy, 711
  hybrid, 733
  urethane, 721
Adipic acid, 853
Aeration, 36
Agrochemicals
  discovery, 1180–1183
  formulation, 1176–1179
  manufacture, 1174–1176
  nomenclature, 1152–1153
Air-free explosions, 137
Air separation, 449
  cryogenic, 442
Air sterilization, 931
Alanin, 953
Alcohols, for gasoline, 495, 507
Alizarin, 888
Alkyd resin, 238
Alkylation, for gasoline, 495, 506
Alkylbenzenes, 853
Allophanate formation, 723
Allyl alcohol, 836
Allyl chloride, 836
Alpha olefins, 826
Aluminum alkyls, 830
Aluminum chloride, 223
Aluminum trihydrate, 251
Amines, aliphatic, 1109–1117
Amino acids, 953
Aminoanthraquinone, 882, 906
7-Aminocephalosporanic acid, 977
6-Aminopenicillanic acid, 974
Ammonia, 1069–1072, 1084–1090
  in fertilizerz, 372–377
  liquid viscosity, 1086
  physical properties, 1072
  plant licensors, 1089, 1090
  price, 1090
  solubility, 1086
  storage & transport, 1090
  synthesis (see Ammonia synthesis)
  uses, 1090
Ammonia synthesis, 1084–1089
  argon recovery, 1088
  catalysts, 1088
  converter details, 1086
  gas composition, 1085
  heat of reaction, 1085
  pressure selection, 1085
  reaction rate, 1085
Ammoniacal copper arsenite, 247
Ammonium nitrate, 1099–1102
  and fuel oil (ANFO), 1101
  crystalline forms, 1101
  fertilizer, 375
  granulation & prilling, 1100
  phase stabilized (PSAN), 1213
  processes, 1100
  safety, 1101

Ammonium nitrate *continued*
uses, 1101–1102
Ammonium nitrate-urea, 375
Ammonium perchlorate, 1213
Ammonium phosphate fertilizer
consumption, 379, 381
diammonium, 386
monoammonium, 388
Ammonium sulfate fertilizer, 372, 375
Amylase, 962
Anaerobic digestion, 37
Analgesics, 992
Analysis of variance, 99
ANFO, 1190, 1193, 1197
Angiosperms, 208
Aniline
physical properties, 1135
price, 1137
processes, 854, 1135–1137
producers, 1137
uses, 1137
Anisole, 889
Anthracene (from wood), 247
Anthraquinone, 880, 884, 909
Anthraquinone sulfonic acid, 882–886
Anti-allergy agents, 993
Antibacterials, 993
Antibiotics, 973, 993
table, 980
Anticonvulsants, 995
Antidegradants (rubber), 604
Antidepressants, 995
Antihistamines, 995
Antihypertensives, 996
Anti-inflammatories, 997
Antineoplastics, 998
Antipsychotics, 998
Anti-ulcer agents, 998
Antiviral agents, 999
Apatite, 350
API, 485
Arabinogalactan, 212
Arabinose, 254
Aramides, 646
Aramid fibers, 738, 784
Argon, 448, 449
from NH_3 manufacture, 1088
Aromatics for gasoline, 495, 499
Ascorbic acid, 959
Aspartic acid, 953
Aspen, 208, 243
Aspirin, 852, 992
Associated (natural) gas, 512
Atactic polymer, 627
Autoignition temperature, 135, 141
Automotive coatings, 1059
Auxochrome, 891
Avermectin, 917, 983, 1150
Bacillus thuringiensis, 1141, 1150, 1181
Badger-Isothermal process for phosphoric acid, 358
Bagasse, 215, 316, 341
Bakelite, 624
Balsa, 208
Bamboo, 214
Basic blue dyes, 869, 902, 906
Basic Brown, 1, 868
Basic Green, 4, 902
Basic Orange, 2, 894
Basic orange dyes, 894, 905
Basic red dyes, 869, 902
Basic violet dyes, 869, 902–905
Basic yellow dyes, 869, 904
Basswood, 208, 263
Beet sugar, 331
Benzanthrone, 889, 906, 909
Benzene, 495, 849
Benzidine, 882
Benzoic acid, 857
p-Benzoquinone, 888
Benzyl alcohol, 857
Bergius-Rheinau process, 255
Bhopal gas release, 168
Bicomponent fibers, 774
Binders, propellant, 1215
Biobleaching, 227
Biochemical oxygen demand, 73
Biological oxygen demand, 73
definition, 74
pulp mill effluent, 234
Biopharmaceuticals, 983
Biopulping, 227
Biotechnology in agrochemicals, 1179
Bipropellant, 1219
Birch, 209, 263
Bisphenol A, 851
diglycidyl ether, 713
resin, 714
Bisphenol F
diglycidyl ether, 714
Bisulfite pulping, 221
Bitumen, 481
Biuret
formation, 723, 1103
of hexamethylene diisocyanate, 725
Black liquor, 220, 221, 266, 270
Black locust, 207
Black powder, 1191
Black walnut, 207
Blanco Directo process, 329
Blast wave, 160
Bleaches, 438
BLEVEs (boiling liquid expanding vapor explosion), 159
Block valves, 127
Blow molding, 694
Boiler explosion, 157
Boiling water reactor, 1266
Breaking length, 736
Bright stock, 484
Brimstone, 464
Brine chemicals, 437
Brix scale, 323
Bromine, 436, 437
Bronchodilators, 993
Bucherer reaction, 891
Buffer zones,
residential areas, 120
Builders, detergent
carbonates, 1039
phosphates, 1039
sodium nitroacetate, 1040
zeolites, 1040
Burning velocity, 135
Butadiene, 612, 846
1,4-Butanediol, 808
Butanes, 489, 495, 506, 838
Butanol, 951
t-Butanol
for gasoline, 507, 841
manufacture, 841
Butene-1, 842
Butyl alcohols, 834
Butylamine, 846
Butylenes, 841
Butyl phenol, 845
Butyl rubber, 615, 845
Butyraldehydes, 834

Calcium chloride, 437, 438
Calcium cyanamide, 1069
Calcium cyanide, 1070
Calcium disulfide, 461
Calcium sulfate
sulfuric acid production, 473
Calutron, 1245
Camphene, 267
Candelilla wax, 282
Candida tropicalis, 270
Cane sugar
manufacture, 316
refining, 325
Capability assessment, 98
Caprolactam, 853
Carbamate agrochemicals, 1161–1163
Carbazole, 912
Carbon dioxide, 452
Carbon disulfide, 809
Carbon fibers, 790
Carbonization, wood, 259
Carbon tetrachloride, 807
Carbonyl disulfide, 461
Cardiovascular agents, 999

Carotenoids, 282
Cascara, 270
Casinghead gas, 512
Catalytic cracking, 494, 501
Catalytic reforming, 492, 499
Catastrophic failure, 125
Catechol, 269
Caustic,
  distribution, 422
  electrolytic cells, 432
  production, 421
Cedar, 246
Cellophane, 669
Celluloid, 623
Cellulose
  derivatives, 208
  structural formula, 211, 668
  wood, 208–270
Cellulose acetate, 752
Cellulose esters, 669
Cellulose ethers, 669
Cellulose nitrate, 623
Cellulose triacetate, 754
Cephalosporins, 975
Chain transfer, 635
  in polymerization, 662, 666
Chamber process, 467
Charcoal, wood, 259–264
Chemical explosives, 1187
Chemical industry
  balance of trade, 3
  capital spending, 4
  definition, 1
  investment trends, 4
  manpower costs, 8
Chemical producers, major
  Europe, 12
  Japan, 12
  United States, 11
Chemical releases, 182
  detection, 183
  mitigation, 183
Chemical variants (fibers), 796
Chemimechanical pulp, 216, 218
Chernobyl, 1228
Cherry, 207
Chestnut, 269
Chicago Acid, 897
Chipping, wood, 217, 235, 236
Chlor-alkali
  manufacture, 421–435
  plant flowsheet, 431
  production, 421
  uses, 422
Chlor-alkali cells
  configurations, 425
  diaphragm, 422, 428
  membrane, 423
  mercury, 423
Chlorinated ethanes, 815
Chlorinated ethylenes, 815
Chlorinated polyethylene, 620
Chlorine
  distribution, 422
  electrolytic cells, 429
  flammability limits, 136
  liquefaction, 431
  producers, 434
  production, 421, 433
  supply & demand, 435
Chlorine dioxide, 439
1-Chloroanthraquinone, 886
Chlorobenzene, 855
Chloroform, 807
Chloromethanes, 805
Chloroprene rubber, 616
Chlorosulfonated polyethylene, 620
Chromogen, 871
Chromophore, 891, 902, 906
Chromotropic acid, 887
Citric acid, 917, 940
Claus process, 459, 523
Clavulanic acid, 975
Clean Air Act, 178
1,6 Cleve's Acid, 899
Coal
  analysis, 533, 535
  classification, 531
  combustion, 565–573
  composition, 533
  consumption, 529
  environmental concerns, 538
  gasification, 550
    gasifiers, 553–558
    power generation, 559
    processes, 560–565
    sulfur recovery, 591
    reactions, 551–553
  heating value, 533
  origin, 530
  plastic properties, 534
  production, 528, 592
  pyrolysis, 546–550
  structure, 533
Coal cleaning, 539–542
Coal combustion
  combustors
    fluidized bed, 567, 571
    pulverized coal, 569
    stoker, 568
  efficiency,
  pollution controls, 572
Coal liquefaction
  direct, 573
  indirect, 588
Coal mining, 536
Coal–oil mixtures, 568
Coal preparation, 539
Coal–water mixtures, 568
Coatings adhesion, 1060
Coatings technology
  flow control additives, 1055
  powder coatings, 1056
  radiation cure, 1056
  solvent borne coatings, 1055
  two component coatings, 1055
  UV-curve, 1052
  waterborne coatings, 1052
Coconut oil, 294
Coextrusion, 688
Coke byproducts, 584
Coking, petroleum, 484, 494, 503
Colour Index, 864
Combustion
  dusts, 134
  hazards, 132
  liquids, 134
  products, 133
  solids, 133
  vapors or gases, 134
Community Right-to-Know, 178, 199
Compreg, 242
Compression and transfer molding, 692
Conjugate fibers, 774
Construction paper, 214, 229
Contact process, 468
Continuous filament yarn, 736
Control limits, 92
Copolymer
  block, 667
  definition, 627
  ethylene-vinyl acetate, 640
  styrene-acrylonitrile, 642
  acrylonitrile-butadiene-styrene, 642
Copolymerization, 663
  reactivity ratios, 664
Copper smelter
  sulfuric acid production, 471
Cottonseed oil, 302
Cottonwood, 208, 243
Coupling, 892, 900
Cracking, petroleum, 494, 501
Creosote, wood, 246, 260, 263
Cresol, 224, 247, 260
p-Cresol, 900
Crocein Acid, 884
Cross-linking in coatings, 1059
Crude oil, 480
  analysis, 488, 493
  desalting, 489, 496
  distillation, 484, 490, 497
  production, 481
  products, 484
  refining, 483
  reserves, 481
  sweet/sour, 488
Cryogenic gases, 444

Cryogenic oxidizers, 1219
Crystallinity, polymer, 631, 636
  crystallites, 631
Crystallinity, rubber, 602
Cumene, 834
  for phenol, 850
Cuprammonium cellulose, 748
Curatives, rubber, 603
Curing agents, 658
CUSUM, 97
Cyanocobalamin, 961
Cyanogen chloride, 889
Cyanuric chloride, 889, 896
Cyclohexane, 852
Cyclopentadiene, 847
Cyclosporin A, 979
Cypress, 246

2,4 D, 1143, 1152, 1155
DDT, 1143, 1148, 1151
Debarking, 216, 235, 266
Deflagration pressure, 151
Dehydration, natural gas, 521
Deinking, paper, 228
Delustrant, fiber, 796
Denier, 736
Descriptive statistics, 87
Detergent additives, 1040
Detergent builders. (*See* Builders)
Detonation
  in chemical plants, 151
  shock, pressure, velocity, 1191, 1196
Deuterium, 1246
Dextrose, 338
Dianisidine, 882, 897
Diaphragm cell
  chlor-alkali, 422
  current voltage, curve, 426
Diazotozation, 892, 900
Dibenzanthrone, 889, 908
Dichlorobenzenes, 856
Dichloroethane, 817
Dicyandiamide, 718
DIERS, 126
Diesel fuel, 484, 486, 492
m-Diethylamino phenol, 905
Diethylene glycol, 813
Diethyl sulfate, 830
Dimensional stability, wood, 242
N-N-Dimethylaniline, 902
Dimethyl sulfide, 270
Dimethyl sulfoxide, 270
Dimethyl terephthalate, 859
DIN, 485
Dinitrophenol agrochemicals, 1170
Dipentene, 267
Direct Black 22, 870
Direct Black 80, 899
Direct blue dyes, 877, 913
Direct Orange 26, 870
Direct Orange 61, 899
Direct Red 83, 898
Direct yellow dyes, 898
Disperse Blue 27, 871
Disperse blue dyes, 871, 895, 907
Disperse Orange 25, 871
Disperse red dyes, 871, 895, 907
Disperse Violet 28, 885
Disperse Yellow 42, 892
Disperse yellow dyes, 871, 892, 900
Distillation, wood
  destructive, 259, 262, 265, 268
  dry, 259
Diuretics, 1000
Dodecene, 835
Dope dyeing, 744, 758, 796
Douglas Fir, 208
Drawing (fibers), 762, 766
  draw texturing, 769
Drug delivery systems, 1009
Drugs, 987
Drying oils, 308
Dry spinning. (*See* Spinning)
Dust
  combustion, 124
  explosions, 123
Dyeing methods, 874
Dyeing textile fibers
  acetate, 758, 866, 870, 873
  acrylic, 866, 873
  cotton, 864, 867, 872
  dimethyl nylon 3, 763
  dope dyeing, 758
  flax, 865
  natural, 864
  nylon, 866, 870
  orlon, 866
  polyester, 866, 873
  polyolefin, 866, 873
  rayon, 865, 872
  regenerated, 865
  silk, 865, 868, 873
  Spandex, 867, 873
  synthetic, 866
  viscose, 747, 758, 867
  wool, 865–873
Dye intermediates, processes for
  addition, 889
  alkaline fusion, 886
  alkylation, 889
  amination, 882
  carboxylation, 889
  condensation, 889
  cyanoethylation, 870
  Friedel–Crafts reaction, 878
  halogenation, 878, 884
  nitration, 878
  oxidation, 888
  reduction, 878, 880
  Sandmeyer reaction, 890
  sulfonation, 878, 883
Dyes
  acid, 864, 867, 873, 906, 923
  acridine, 905
  anthraquinone, 906
  application of, 873
  azine, 905
  azo, 891
  azoic, 864
  basic, 864, 868, 873, 913
  classification, 867
  diphenylmethane, 904
  direct, 864, 869, 873, 913
  disazo, 896
  disperse, 864, 870, 873, 907, 913
  indigoid, 910
  manufacture, 900
  methine, 869
  monoazo, 894
  mordant, 864, 876, 913
  nitro, 892
  non-textile uses, 870
  oxazine, 870, 905
  phthalocyanine, 870, 912
  polyazo, 899
  production and sales, 913
  reactive, 864, 871, 873, 913
  solvent, 864, 913
  stilbene, 870, 899
  sulfur, 864, 872, 911
  thiazine, 905
  thiazole, 870
  thioindigoid, 910
  triphenylmethane, 902
  vat, 864, 872, 908, 913
  vinyl sulfone, 872
  xanthene, 904
Dynamite, 1192, 1199, 1204

Eastman chemicals from coal, 587
Economy, United States, 3
Edible fats and oils
  bleaching, 300
  degumming/refining, 299
  fatty acids, 278
  hydrogenation, 300
  products, 302
  refining, 297
  solid fat, 297
Elaidic acid, 275
Elastomer
  definition, 625
  microstructure, 600
Elastomeric fibers, 776
Electric arc, 1069
Emergency organizations, 184, 191, 199

Emergency response program
agencies, 203
communications, 189, 191, 196, 201
community plans, 197
control center, 188
drills, 195, 199, 202
emergency preparedness team, 184
evacuation, 187
fire brigade, 184, 194
laws and regulations, 178, 200, 202
management, 184, 191, 194
media relations, 189, 191
response plan, 178, 185, 189, 201
Emulsifiers, 304
Emulsion polymerization, 608
Emulsion styrene-butadiene rubber, 607
Enantiomorphic selection synthesis, 13
Energy consumption, world, 481
Energy supply and use
fossil fuels, 1236
per capita use, 1234
uranium, 1236
U.S. flowsheet, 1235
world flowsheet, 1234
Enzymes
industrial catalysts, 962
Epichlorohydrin, 712, 836
Episodic risk, 165
Epoxy adhesives, 711
epoxy reactions, 716
resin manufacture, 713
Epoxy phenol novolac, 714
Epoxy resins, 657
for adhesives, 713
Erythorbic acid, 947
Estrogen, 1000
ETBE, 495, 508
Ethanol
by fermentation, 948
for gasoline, 495, 507
from wood, 254, 257, 262, 270
motor fuel, 254
synthesis, 819
Ethanolamines, 814
Ethylbenzene, 821
Ethyl chloride, 817, 898
Ethylene
as raw material, 809
copolymers, 830
from steam cracking process, 838
manufacture, 495
products from, 810
uses, 830
Ethylenediamine
physical properties, 1115
price, 1117
processes, 1115
producers, 1116
uses, 1116
Ethylene dichloride, 817
Ethylene glycols, 813
Ethylene oligomers, 826
Ethylene/propylene rubber, 614
Explosion
BLEVEs, 150
damage, enclosed structures, 121
dust, 123
in absence of air, 137
mists, 141
physical, 157
pressure, 151
static electricity, 146
unconfined vapor clouds, 156
violence, 153
Expansion joints, selection, 129
Explosion pulping, 236
Explosives
applications, 1187
chemistry, 1189
classification, 1198
composite, 1190
consumption, 1207
emulsion, 1205
high, 1198
historical, 1191
liquid oxygen, 1190
low, 1198
manufacture, 1199
molecular, 1189
primary, 1196
properties, table, 1196
safety, 1209
secondary, 1196
slurry, 1205
structure, 1199
tertiary, 1197
usage, table, 1208
water gel, 1193, 1197, 1205
Extractives, wood, 212, 264
Extruder dies, 686
Extrusion (fibers). (*See* Spinning)
Exxon donor solvent process, 574

Fat
refining, 1021
splitting, 1021
substitutes, 305
Fats and oils. (*See* also Edible fats and oils)
analytical methods, 310
antioxidants, 308
chemical structure, 374
consumption, 287
lipolysis, 308
oxidation, 308
sources, 285
trade, 312
Fatty acids, 274–278, 1020–1030
alternative names, 276
characteristics, 276
crystallization, 1029
distillation, 1020, 1026
enzymatic, 1026
epoxy, 308
fractionation, 1020, 1027
hydrogenation, 1021
hydrophilization, 1029
in edible oils, 278
in industrial oils, 279
methyl ester process, 1026
neutralization, 1015, 1020
polymorphism, 279
pressing, 1029
production, 1020
refining, 1020
separation, 1029
unsaturation, 1030
FD&C Red 40, 894
FD&C Yellow 5, 895
Fermentation
fundamentals, 918
microbial kinetics, 922
microorganisms, 918
oxygen transfer products
amino acids, 953
antibiotics, 973
enzymes, 962
organic acids, 937
organic solvents, 948
vitamins, 959
Fermenters, 926
Fertilizers
classification, 368
consumption, 370
industry overview, 369
nitrogen
mineral nitrogen, 371
natural organic, 371
synthetic ammonia, 372–377
mixed (see Mixed fertilizers)
phosphate, 377–388, 477
potassium, 388–393
Fiber. (*See* by type)
Fiberboard, 208, 234
Fiber variants, 793
Filament winding, 656
Filament yarn, 736
Fillers, rubber, 603
Fir, 216
Fire
definition, 133
hazards, 143
products, 133
triangle, 134

Fire and Explosion Index, 127
Fire prevention, 186
Fisher's aldehyde, 869
Fisher's base, 869
Fish oils, 295
Fission, nuclear, 1238
Fission products, 1241
FL 506, 979
Flakeboard, 208, 245
Flame
  arresters, 153
  quenching, 136
Flammability
  limits, effect of
    inert gases, 138
    pressure, 137
    temperature, 136
    oxygen enrichment, 136
  heterogeneous mixtures, 137
  mists and foams, 140
  with chlorine oxidant, 136
  with nitric oxide, 136
Flammable gases, 442
Flash point, 135
  mists, 141
Flexible joints, 129
Flixborough disaster, 120, 167
Fluorine, phosphate rock, 364
F–N curves; risk studies, 166
Formaldehyde, 804
Fourdrinier papermachine, 230, 236
Frasch process, 459
Friedel–Crafts reaction, 878
Fructose, 340
  high fructose syrup, 339
Fuel elements, nuclear, 1266
Fuel, rocket
  Aerozine-50, 1222
  cryogenic, 1222
  gelled, 1223
  liquid, 1219
  metallic, 1215
  solid, 1215
  storable, 1222
  UDMH, 1222
Furfural, 208, 269
Furfural alcohol, 269
Fusion, 1239

G-Acid, 877, 884, 896
Galactose, 254
Gamma Acid, 893, 899
Gas compression, 523
Gaseous diffusion process, 1252
Gasification, coal, 550–565
  gasifier types, 553
  processes
    GEGAS-D, 561
    GE moving bed, 560
    Koppers Totzek, 564
    KWB, 564
    Texaco, 562
    rotary kiln, 560
    Westinghouse, 560
Gasification, wood, 259
Gasohol, 254, 590
Gasoline, 484–490
  ethers for, 495, 507
  Fisher Tropsh synthesis, 588
  from methanol (Mobil), 592
Gaussian curve, 89
Gel coats, 656
Gel effect, 671
Gelled fuels, 1223
Gelled oxidizers, 1224
Gelled propellants, 1223
Gel spinning, 782
Gibberellic acid, 1150
Glass fibers, 737, 789
Glass transition temperature, 602, 632, 676
Glucoamylase, 963
Glucomannan, 212
Gluconic acid, 945
Glucono-δ-lactone, 945–946
Glucose, 213, 254, 256, 337
  isomerase, 339, 965
  syrups, composition, 337
Glutamic acid, 955
Glycerin, 836, 905, 1021
Glycol ethers, 815
Graphite fibers, 790
Gross National Product, 2
Groundwood pulp, 216
Gymnosperms, 208
Gypsum, 466

Haber process, 1069
H-Acid, 887, 893, 896
Hardboard, 234
Hazard
  evaluation, 163, 164
  identification, 180
  mitigation, 180, 182
Hazardous materials
  Federal laws, 33
Hazardous waste
  cleanup, 65
  definition, 60
  landfill, 65
  management, 53, 61
  treatment and disposal, 31, 63
Hazardous waste treatment
  costs, 64
  technologies, 58, 64
HAZOP, 120, 164
H-Coal, 574
Heating oil, 484, 492
Heat setting, 766
Heavy metals
  from liquid waste, 52
Helium, 452
Hemicellulose, 212, 223, 255
Hemlock, 269
Hemp, 215
Heptenes, 843
Hexahydrophthalic acid, 715
Hexamethylenediamine, 853
Hexamine, 1117
  processes, 1118
  properties, 1118
  uses, 1120
High fructose syrup, 339
High performance fibers, 782
HMX, 1200, 1213
HNS, 1200
Hoffman process, 1123
Holocellulose, 255
Hottenroth number, 745
Howard process, 269
Hydrazine, 1120–1128
  handling, 1126
  processes, 1122
  properties, 1121
  safety, 1121
  uses, 1127
Hydrocracking, 494, 5004
Hydrogen
  applications, 450
  ortho and para, 449
  ultra-pure, 450. (*See* also Hydrogen production)
Hydrogen cyanide,
  physical properties, 1128
  price, 1135
  processes, 809, 831, 1128
  producers, 1134
  safety, 1128
  uses, 1134
Hydrogen production
  carbon dioxide removal, 1081
  carbon monoxide shift, 552, 1080
  chlor-alkali cells, 429
  coal/coke gasification, 550, 1079
  electrolysis, 1080
  naphtha treatment, 1075
  natural gas treatment, 1074
  oxidation of CO
  partial combustion, 1078
  petroleum gases, 1080
  processes, 1073, 1075
  raw materials, 1074
  reforming, 1075
  synthesis gas, 1074
Hydrogen sulfide, 458
Hydrophile, 1031
Hydrophobe, 1031
Hydrotreating, 491, 498

Hypocholesteremics, 1001
Hypochlorite bleaching, 439
Hypoglycemics, 1001

Ignition
  delay, 142
  energy, 144
  sources, 141
  spontaneous, 143
  static electricity, 149
Impreg, 242
Indamine, 906
Indigo, 910
Indoxyl, 910
Industrial experimentation, 98
Industrial gases
  cryogenic air separation, 442
  transport, 445. (*See* also individual gases)
Inert gases, 442
  data credibility, 140
  effect on flammability limits, 138
  hazards, 140
Injection molding, 688
Inks
  gravure, 1061
  letterpress, 1058
  lithographic, 1058
  screen, 1062
Insect hormones, 1168
Insect repellants, 1172
Insulation board, 234
Isoascorbic acid, 947
Isobutane, 841
Isobutylene, 843
Isocyanates, 651
Isophthalic acid, 859
Isoprene, 607, 847
Isopropyl alcohol, 832
Isotactic polymers, 627, 637, 674
Itaconic acid, 944

J-Acid, 887, 893
J-Acid urea, 898
Jet fuel, 484, 486, 492
Jojoba, 282
Jute, 215

Kamyr digester, 219
Kerogen, 481
Kerosine, 484, 492
2-Ketogluconic acid, 946
Kraft paper, 241
Kraft process, 215, 265

Lacquers, 1066
Lactic acid, wood, 270
Lakes, 1064
Lards, 295
Latex, 604, 610, 672
Lean gas, 510
Le Chatelier's equation, 135
Lecithin, 305
Lewis acid, 720
Lignases, 227
Lignin, 208, 212
Linear primary alcohols, 826
Linoleic acid, 267
Linseed oil, 238
Lipids, 273
Liquefied natural gas, 454, 519
Liquefied petroleum gas, 484, 486
Liquid crystalline polymers, 770, 784
Living polymers, 667
LMFBR, 1269
Lubricants, 484
Lysine, 955

Magnefite pulp, 216
Magnesium hydroxide, 437
Maleic anhydride
  from benzene, 853
  from butane, 840
Mannose, 254, 270
Manufactured gas, 513
Maple, 263
Mechanical pulp, 216
Mechanistic model building, 112
Median, 87
Melamine, 1107
  formation, 1108
  price, 1109
  processes, 1108
  producers, 1109
  uses, 1109
Melt blown fibers, 781
Melt spinning. (*See* Spinning)
Membrane cell, 423
Membrane separation, wastewater, 42
Mercury
  in smelter gases, 472
Mercury cell, 423
Methane, chemicals from, 801
Methanol
  for gasoline, 495, 507
  from hydrocarbons, 803
  from syn gas, 590
  from wood, 259
Methyl acetate
  Eastman process, 592
  from wood, 263
Methylamines, 1109
  processes, 1110
  properties, 1110
  uses, 1115
Methyl-*t*-butyl ether, 805, 832, 843
Methyl chloride, 805
Methyl diisocyanate, 854
Methylene chloride, 807
Methyl ester, 1014
Methyl ethyl ketone, 840, 842
Methylmethacrylate, 804
MHD, 565
Micelles, 672
Michler's ketone, 903
Microbial kinetics, 922
Microdenier fibers, 793
Mists
  combustion, 134
  flammability, 140, 142
MITI, 485
Mixed fertilizers, 394
Mixture experiments
  centroid design, 109
  Scheffe model, 111
  simplex lattice, 109
Mobile incinerators, 42
Modacrylic fibers, 737, 776
Molasses, 316, 340
Monarch process, 475
Monobactam, 979
Monopropellant, 1223
Mordant Black 17, 895
Mordant Brown 4, 895
MTBE, 495, 508
Multiple regression, 89

Naphthalene, 859
Naphthalenesulfonic acid, 883
Naphthionic acid, 884, 891
Naphthol, 884, 887, 893
1-Naphthylamine, 884, 887
2-Naphthylamine, 891, 896
Natural gas
  absorption, 524
  composition, 511
  heating value, 511
  liquids recovery, 524
  petrochemical products, 515
  pipeline quality, 512
  reserves, 518
  U.S. production, 514
  world reserves, 519
Natural gas hydrate, 521
Natural gas liquids, 514
Naval stores, 263
Neat soap, 1015
Neutral fat saponification, 1015
Newsprint, 214, 229
Nitric acid, 1090
  ammonia oxidation
    catalysts, 1095

Nitric acid *continued*
chemistry, 1091
equilibrium constants, 1092
concentration, 1096
physical properties, 1091
pollution abatement, 1098
processes, 1093
production rates, 1099
purification, 1098
stabilizers, 1098
uses, 1099
Nitrides, 1069
Nitrile rubber, 616
Nitrobenzene, 854
Nitrocellulose, 1212
Nitro Delta Acid, 889
Nitrogen, 442
applications, 447
by PSA process, 446
occurrence, 447
production, 447
uses, 447
Nitrogen fertilizers, 370
Nitrogen fixation, 1069
Nitrogen oxides, 456, 467, 1069
Nitrogen products, 1068–1140
Nitroglycerine, 1189, 1196, 1203
Noguchi process, 264
Nonene, 835
Normal distribution, 89
Nuclear applications, 1270–1274
Nuclear industry, 1226
Nuclear processes, 1236
Nuclear reactor materials processing, 1244
Nuclear reactors
characteristics, 1265
containment, 1268
fuel elements, 1266
Oconee Nuclear Station, 1229
types, 1266
uranium requirements, 1265
Nuclear safety, 1230
Nuclide production, 1240
Nylon, 646
Nylon fibers, 737, 758
nylon 6, 750
nylon 66, 759
other, 763

Octane number, 491, 500
Octenes, 843
Oilseeds
screw pressing, 293
solvent extraction, 288, 290
Oil shale, 481
Olefin fibers, 737
Oleic acid, 275
Olein, 275, 1030
Oleoresins, 265
Oleum, 467
Oligomer, 627
Olive oil, 294
Organosolv pulping, 223
Overpressure
detonations, 152
explosion effects, 154
Oxidizers, gases, 442
Oxidizers, propellant
ammonium nitrate, 1213
ammonium perchlorate, 1213
chlorine pentafluoride, 1221
chlorine trifluoride, 1222
cryogenic, 1221
dinitrogen tetroxide, 1221
fluorine, 1221
gelled, 1224
HDA, 1222
hydrazinium nitrate, 1213
hydrazinium perchlorate, 1213
hydrogen peroxide, 1222
hydroxylammonium nitrate, 1213
hydroxylammonium perchlorate, 1213
liquid oxygen, 1221
MON, 1222
nitrogen tetroxide, 1221
Oxo chemicals, 805, 834
Oxygen, 447
Ozonation, 43
Ozone bleaching, 4390

Paint and coatings,
environmental regulations, 1051
historical, 1056
markets, 1050
production, 1051
technology change, 1052
Palm oil, 295
Paper
consumption, 214
converting, 232
finishing, 232
opacity, 229
production, 215
recycling, 228
strength, 225
Paperboard, 214
Papermaking, 213, 229
Paraformaldehyde, 804
Parenchyma, wood, 209
Parenteral preparations, 1008
Partially oriented yarn, 768
Particleboard, 208, 238
PCB soils
decontamination, 46
Peanut oil, 303
Penicillin amidase, 963, 975
Penicillins, 917, 973, 1006
Pentaerythritol tetranitrate, 1203
Peracidox process, 474
Perchloroethylene, 819
Perilla oil, 238
Pesticides
production, 1144
regulation, 1145
PETN, 1203
Petrochemical feedstocks
from coal, 584
Petrochemicals, 487, 495
Petroleum, 480
Pharmaceutical industry, 987
production, 1006
quality assurance, 1010
Phenol, 260, 850, 880, 889
Phenol-formaldehyde resins, 624, 653
Phenolic resins, 653
novolacs, 654
resoles, 653
Phenoxy resins, 657
Phenylglycine, 910
Pheromones, 1171
Phillips explosion, 1170
Phosgene, 903
Phosphate fertilizers
consumption, 370, 379
from mineral phosphates, 378
natural organic, 378
production, 381
Phosphate minerals, 350
Phosphate rock
analysis, 350
exporters, 349
fertilizers, 378
occurrence, 347
resources, 348
Phosphates, industrial, 354
Phosphogypsum, 363, 472
Phospholipids, 273, 304
Phosphoric acid
dihydrate processes, 355
environmental aspects, 366
hemihydrate process, 358
production, 355
purified, 365
superphosphoric acid, 362
wet process, 360
Phosphorus, 352
Phosphorus agrochemicals, 1159
Phthalic anhydride, 888, 904, 909, 913
by naphthalene oxidation, 860
from *o*-xylene, 859
Phthalocyanines, 912
Physical explosions, 157
Physical explosives, 1187
Picric acid, 892

Pigments
aluminum, 1066
extender, 1063
hiding power, 1062
inorganic, 1063
organic, 1063–1066
pearlescent, 1066
Pigment Violet 1, 905
Pigment Yellow 3, 896
Pilling (in fibers), 796
Pimaric acid, 267
Pine, 208, 264
Pinenes, 265
Pine oil, 265
Piping supports, 127
Pitch, wood, 263, 266
Plant foods. (*See* Fertilizers)
Plastic coatings, 1059
Plastic foams, 699
environmental issues, 704
Plasticized polyvinylchloride, 619
Plasticizers, 603, 630, 638, 640
Plastics extruders
computer simulation, 685
dual channel screws, 682
foam, 701
grooved barrel, 683
single screw, 677
theory, 680
twin screw, 688
two-stage screw, 681
Plastics fabrication, 677
Plastics manufacture
environmental concerns, 672, 674, 702
Plastics molding
blow, 694
compression and transfer, 692
injection, 688
Plywood, 208
Pollution abatement, 70
Pollution prevention, 15
Polyacrylonitrile, 643
Polyamide polyamine synthesis, 718
Polybenzeneimidazole fibers, 738, 787
Polybutadiene production, 612
Polybutenes, 845
Polycarbonates, 648
Polydispersity index, 629
Polyester fibers, 737, 764
Polyesters
polybutylene terephthalate, 645
polyethylene terephthalate, 644
unsaturated, 654, 693
Polyetherether ketones, 649
Polyethers, 668
Polyethylene
high density, 635, 810
linear low density, 635, 810
low density, 635, 810
Polyethylene fibers, 782
Polyethylene glycol, 814
Polyethylene oxide, 668
Polyethylene oxybenzoate fibers, 765
Polyethylene terephthalate fibers, 764
Polyimides, 647
Polyisoprene, 606
Polymer gasoline, 495
Dimersol polymerization process, 505
Polymerization catalysts
Phillips, 636
Ziegler–Natta, 636, 640
Polymerization mechanisms
addition, 660
free radical, 661
ionic, 664
ring opening, 667
condensation, 659
epoxy cure
chain growth, 716
step growth, 720
Polymerization methods, 670
Polymer networks, elastomers, 601
Polymer nomenclature, 626
Polymers
molecular weight
number average, 627
weight average, 629
thermoplastic, 630
thermoset, 630
structure, 626
networks in elastomers, 601
synthesis, 659
Polymethylene, 627
Polymethyl methacrylate, 643
Polyolefin fibers, 778
Polyolefin orientation
uniaxial, 638, 644
biaxial, 638, 644, 677, 687
Polyolefins, 634
Unipol process, 674
Polyphenylene sulfide fibers, 792
Polyphosphates
in fluid fertilizers, 401
Polypropylene, 636, 674, 831
Polypropylene fibers, 778
Polystyrene, 640
rubber, 641
Polytetrafluoroethylene fibers, 788
Polyurethane, 651, 668
foams, 652
rubber, 618
Polyvinyl acetate, 639, 664
Polyvinyl alcohol, 639, 669
fibers, 775
Polyvinyl butyral, 640, 669
Polyvinyl chloride, 638
fibers, 775
Polyvinylidene chloride, 639, 664
Potassium fertilizers
compounds, 388
consumption, 370
manufacture, 388
mineral sources, 389
Prayon process, 357
Prescription drugs, 987
Pressure relief systems, 126
Pressurized water reactor, 1265
Printing dyes, 876
Printing paper, 229
Process control, 97
Process safety, 181
Propellant. (*See* Rocket propellants)
Propionaldehyde, 830
Propyl alcohol, 830
Propylene, 830
for gasoline, 495, 505
oligomers, 835
Propylene glycols, 832
Propylene oxide, 831, 841
production, 10, 13
Proteases, 967
Pulp
bleaching, 220, 234
production, 215
screening, 225
strength, 225, 229
Pulp digester, 219
Pulping, 213
processes, table, 216
pollution control, 223
Pultrusion, 656, 693
Pumps
explosions, 161
hazardous service, 129
Purex process, 1255
Pyrethroid insecticides, 1152, 1165
Pyridine, 247
Pyrite roaster, 471
Pyrogallol, 269
Pyroligneous acid, 259, 263
Pyrolysis
coal, 546
wood, 259

Quality
total quality, 114
Quantitative risk assessment, 120, 165
Quebracho, 269
Quinizarin, 907

R-Acid, 884, 896
Radiation processing, 1270
Radiation sources, 1272
Radioactive waste management, 1258
Radioisotope tracers, 1274
Rapeseed/canola oil, 303
Raw gas, 510
Rayon
  fibers, 737, 741–752
  yarn, 669
RDX, 1200, 1213
Reaction injection molding, 652, 691
Reactive chemicals, 122
Reactive Orange 13, 896
Recycling, reuse, 7
  waste reduction, 66
Redwood, 246
Refiner pulp, 216
Reflected pressure, 152
Remelt sugars, 329
Rennin, 969
Reppe process, 729
Resin
  definition, 624
  production, 623
Resin acids, 266
Resorcinol, 887, 904
Rheology, polymer, 668
Rhone Poulenc processes, 357
Riboflavin, 960
Rich gas, 510
Risk, 131
  assessment, 165
  contours, 166
  evaluation, 163
  studies, 166
Rocket fuel. (*See* Fuel, rocket)
Rocket motor, 1210
Rocket propellants
  ballistic properties, 1217
  bipropellant, 1219
  boost, 1212
  double-base, 1212
  gelled, 1223
  hazards, 1218
  liquid, 1210, 1218
  manufacture, 1218
  monopropellant, 1223
  oxidizer, 1213, 1221
  single base, 1212
  solid, 1210
Rockets, 1209
Rodenticides, 1172
Rosin, 208, 229, 266
Rotational molding, 698
Rotenone, 1150
Rubber
  antidegradants, 616
  butyl, 615
  chloroprene, 616
  compounding, 603
  copolymer, 601
  crystallinity, 602
  curatives, 603
  emulsion styrene-butadiene rubber, 607
  ethylene-propylene rubber, 614
  fillers, 603
  latex, 604, 610
  microstructure, 600
  natural, 604
  networks, 601
  nitrile rubber, 616
  plasticizers, 603
  polyurethane rubber, 618
  reinforcing agents, 603
  silicone elastomers, 617
  solution styrene-butadiene rubber, 611

Saccharification, cellulose, 252
Saccharomyces cerevisiae, 270
Safety
  hazards, inert gases, 140
  incidents, 167
  in design and operation, 119
  management, 130
  systems, 162
Safflower oil, 307
Safranine B, 906
Salicyclic acid, 852, 889
Salt dome, 458
Sandmeyer reaction, 890
Saponification, 1013
SASOL, 587
Schaeffer's Acid, 884
Scholler-Tornesch process, 255
SCOT process, 461
Secondary cellulose acetate fibers, 752
Segmented polyurethane fibers, 778
Semichemical pulp, 216, 223, 229
Semiochemicals, 1171
Sewage, fertilizer, 371
Sheet molding compound, 657
Shewhart chart, 92
Silicone elastomers, 617
Silicone resins, 659
Silvichemicals, 252
Single cell protein, 969
Sistosterol, 267
SITE program, 58
Soap
  chemistry, 1012
  definition, 1012
  drying, 1015
  feedstocks, 1013
  finishing, 1019
  lathering, 1015
  manufacture, 1013
  packaging, 1020
  pellets, 1015
  properties, 1014
  solubility, 1014
Soapstock, 1026
Soda ash, 409
Soda pulping, 241
Sodium bicarbonate, 413
Sodium bisulfite, 418
Sodium carbonate. (*See* Soda ash)
Sodium chlorate, 440
Sodium chloride, 408
Sodium hyposulfite, 418
Sodium nitrate, 1213
  fertilizer, 371
Sodium phosphates, 419
Sodium silicate, 420
Sodium sulfate, 414
Sodium sulfides, 416
Sodium sulfite, 417
Sodium thiosulfate, 417
Soil fumigants, 1170
Solution styrene-butadiene rubber, 611
Solvent Green 3, 907
Solvent pulping, 223
Solvent red dyes, 877, 905
Solvent refined coal, 577
Solvent Yellow 1, 894
Sorbitol, 959
Sour gas, 458
Soybean oil, 302
Space shuttle, 1217
Spandex fibers, 738, 778
Spin drawing, 762, 769
Spinning finish, 749
Spinning of fibers (extrusion)
  dry, 756, 773, 775, 787
  dry jet-wet, 786
  gel, 782
  melt, 761, 766, 780
  wet, 745, 773, 775, 778, 785
Spontaneous ignition temperature, 141
Spray drying, detergents, 1041
Spruce, 216
Spunbonded fabric, 781
Spun yarn, 736, 749
SRC-II, 578
Standard deviation, 88
Standard Industrial Classification, 48
Staple fibers, 735
Starch hydrolysis, 337
Static
  charge, phase separation, 146

discharge
development, 149
types, 146
electricity, 145
filling operations, 150
grounding and bonding, 150
ignition source, 144
relaxation time, 147
Statistical design of experiments. (*See* Industrial experimentation)
Statistical inference, 91
Statistical tools
box plot, 86
dot plot, 86
fishbone diagram, 114
flow chart, 114
histogram, 85
run chart, 84
scatter plot, 114
stem and leaf plot, 86
Statistics, definition, 83
Staybwood, 242
Stearin, 274, 1030
Stereochemistry, 627
Stereolithography, 14
Stilbene, 267
Stone groundwood pulp, 216, 236
Storax, 270
Streptomycin, 979, 982, 1150
Stretford process, 461
Styrene, 608, 640, 849
Sugar
deliveries, 336
direct consumption, 329
environmental concerns, 343
manufacture, 316
production value, 336
refining, 325
trade, 341
Sugar cane, 316
Sugars
agglomerated, 331
amorphous, 331
brown, 331
liquid, 330
pulverized, 331
starch, from, 337
Sulfacid process, 474
Sulfanilic acid, 883, 894
Sulfar fibers, 738, 792
Sulfite process, 215, 221, 258
Sulfolane, 224
Sulfreen process, 461
Sulfur
burner, 469
in fuel, 491
markets, 466
mining, 460
molten, 459
production, 464
recovered, 461
solid, 459
Sulfur dioxide, 469
Sulfuric acid
chamber process, 467
contact process, 468
fuming, 466
industry, 466
mist, 472
production, 476
uses, 466
waste, 472
Sulfurous acid, 467
Sulfur trioxide, 473
Sunflower oil, 303
Supercritical water oxidation, 39
Superphosphate fertilizer,
"ordinary", 380
"triple", 383
production data, 381
Superphosphoric acid, 362
Surfactants
alkylamides, 1038
amphoteric, 1038
anionic, 1032
cationic, 1038
linear alkylbenzene, 1032
nonionic, 304, 1037
Sulfonation, 1033
Sweetening natural gas, 523
Syndiotactic polymer, 627
Synthesis gas, 584, 591, 801, 1074
Synthetic detergents, 1031

Tablets and capsules, 1007
Tall oil, 238, 265, 267
Tallows, 295TAME, 495, 508
Tannins, 208, 235, 268
Tar sand, 481
sulfur recovery, 464
Tartaric acid, 895
TATB, 1202
Taxol, 270
Teflon, 788
Telomers, 662
Tempken reaction, 1085
Tenacity, 736
Terephthalic acid, 859
Terpenes, 265, 283
Terpin hydrate, 267
Tetrachloroethylene, 819
Tetracyclines, 977
"Tex", definition, 736
Textile fibers, manufactured, 735
Textured yarns, 767
Thermal radiation, 156
Thermoforming, 696
Thermomechanical pulp, 216, 255
Thermoplastics, 634
Thermosets, 651
Thienamycin, 983
Thioindoxyl, 911
Thiosalicyclic acid, 911
Three Mile Island, 1228, 1268
Titanium dioxide, 744, 796
Tolidene, 897
Toluene
derivatives, 856
petroleum refining, 495
Toluene diisocyanate, 856
M-Toluidine, 892
o-Toluidine, 902
p-Toluidine, 902, 907
Toners, 1064
Transdermal systems, 1009
Triacetate fibers, 737, 754
Trichloroethane, 818
Trichoderma viride
Tricking filter, 34
Triethylene glycol, 814
Triglycerides, 1013
refining, 1014, 1021
splitting, 1020, 1024
Trinitrotoluene, 1189, 1193, 1200
explosion model, 154, 157
Tripelennamine, 1003
Tromsdorff effect. (*See* Gel effect)
Trytophan, 955
Tung oil, 238, 308
Turpentine, 208, 265

Uranine, 905
Uranium enrichment, 1251
Uranium fuel cycle, 1246
Urea
additives, 1107
fertilizer use, 375
processes, 1102
product forms, 1106
Urea agrochemicals, 1163
Urea-ammonium nitrate
fertilizer, 375
Urethane adhesives, 721
UV absorbers
in coatings, 1059

Valves
block isolation, 127
fire safe, 128
Vanillin, 208, 222, 269
Vat blue dyes, 873, 908
Vat Green 1, 908
Vat Orange dyes, 908, 911
Vat red dyes, 911
Vat yellow dyes, 908
Vinal fibers, 775

Vinegar, 937
Vinyl acetate, 826
Vinyl chloride, 10, 818
Vinyl fibers, 775
Vinyl resins, 638
Vinyon fibers, 775
Viscoelasticity, 633, 676
Viscose process, 669
Viscose rayon fibers, 737, 741
Vitamin B_2, 960
Vitamin B_{12}, 961
Vitamin C, 959
Vitamins, fat soluble, 282
VOC regulations, 1052

Waferboard, 239
Wallis amine, 224
Waste minimization
  assessment procedure, 18, 21
  economic benefits, 26
  feasibility analysis, 23
  program implementation, 28
  worksheet, 25
Wastepaper, 228, 235, 258
Wastes, industrial
  EPA and OTA definitions, 20
  management 49
  types, 47
Waste treatment technologies, 37–53
Wastewater
  bioremediation, 45
  from selected industries, 50
  terminology, 72
  treatment levels, 32
  treatment technologies, table, 55
Waterborne coatings, 1052
Water use, industrial, 67
Waxes, 281
Weissenberg effect, 677
Wellman–Lord process, 464
Wet air oxidation, 38
Wet natural gas, 510
Wet spinning. (*See* Spinning of fibers)
Wet strength resins, 229
Whale oil, 282
Willow, 207
Wood pulping processes, 216
Writing paper, 214, 229

X, MR chart, 96
XS chart, 92
Xylan, 212
Xylene, 495, 857
Xylose, 212, 254, 269, 270
Yankee dryer (paper), 232

Zerewitinoff reaction, 722