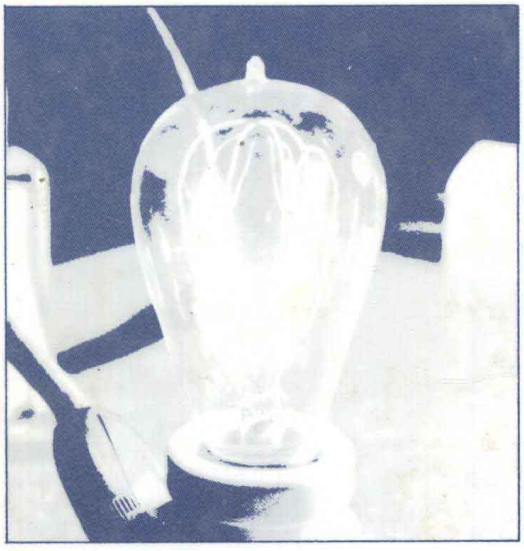
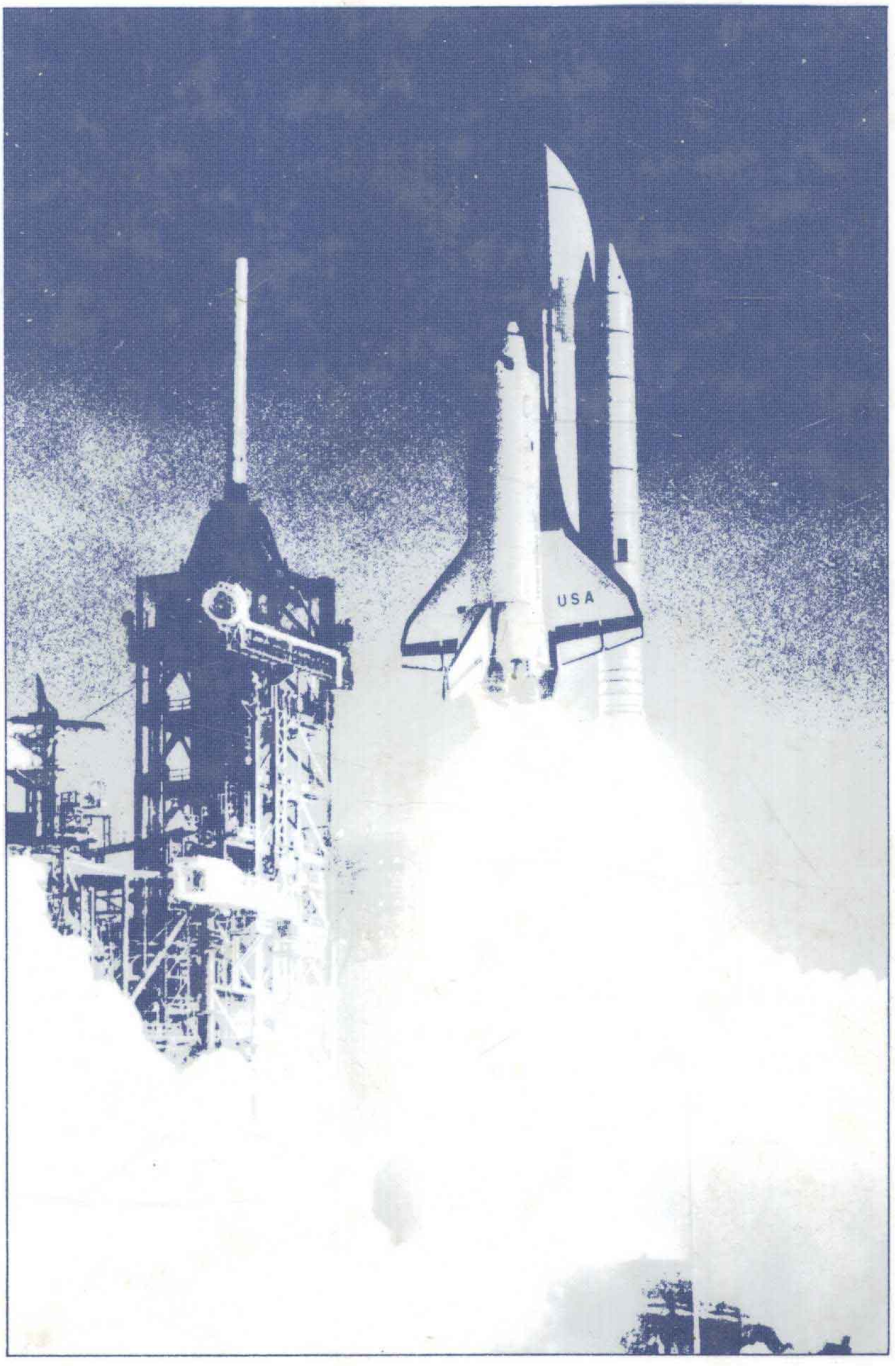
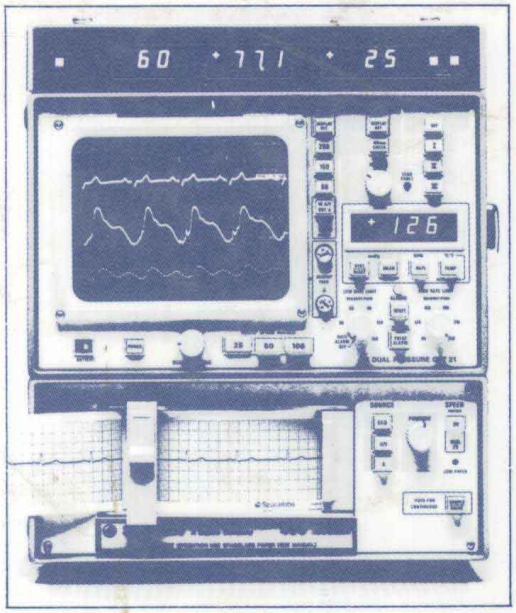
Electronics

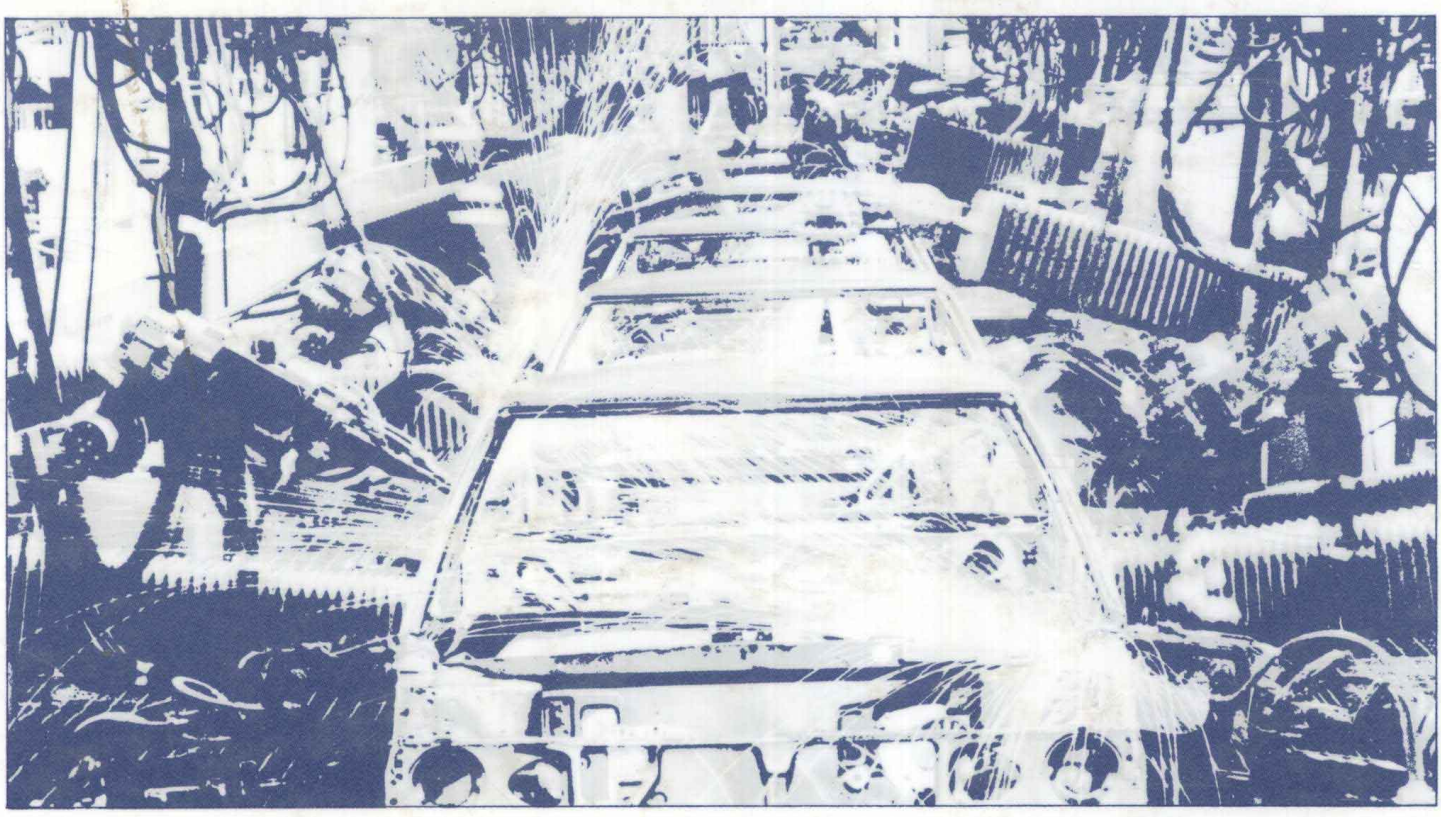
Concepts, Applications, and History

Roger J. Houglum









77F6056

ISBN 0-534-03909-X

ELECTRONICS: Concepts, Applications, and History Second Edition

ROGER J. HOUGLUM
Lane Community College

Breton Publishers
A DIVISION OF WADSWORTH, INC.
Boston, Massachusetts

Breton Publishers
A Division of Wadsworth, Inc.

© 1985 by Wadsworth, Inc., Belmont, California 94002. All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transcribed, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher, Breton Publishers, a division of Wadsworth, Inc., 20 Park Plaza, Boston, Massachusetts 02116.

Library of Congress Cataloging in Publication Data

Houglum, Roger J., 1911-
Electronics: concepts, applications, and history.

Includes index.
1. Electronics. I. Title.
TK7880.H69 1985 621.381 84-9193
ISBN 0-534-03909-X

Printed in the United States of America
1 2 3 4 5 6 7 8 9 — 89 88 87 86 85

ISBN 0-534-03909-X

Electronics: Concepts, Applications, and History, Second Edition was sponsored by George J. Horesta. It was prepared for production by Jean T. Peck. Kathy Leahy served as copy editor. Mary S. Mowrey served as art coordinator and interior designer. The cover designer was Trisha Hanlon. Cover photos: Front, clockwise from top left: Tektronix, Inc.; NASA; by the author. Back, clockwise from top left: Spacelabs, Inc.; Heath Company; Chrysler Corp.

I would like to dedicate this book to the staff members of the Electronics Department of Lane Community College for their constant support and encouragement during the long period required for its research and writing.

Preface

Today electronics touches the life of nearly every American in some manner scores of times in the course of a single day. In the 1980s, the problem of keeping everyone fully informed of new developments in the ever-widening and ever-changing field of electronics can no longer be met entirely by trade journals and technical publications.

For example, computer specialists know little about medical electronics, yet the use of computers and microprocessors has revolutionized the medical profession's approach to patient diagnosis and treatment. And to compound the problem, trained electronics personnel working in the field of medical electronics frequently have little or no understanding of computer and microprocessor operation. What is needed is a book that will serve as a clearinghouse of information for those with a professional, an avocational, or a casual interest in new developments in the field of electronics and the technology of its applications. We hope that this book, with its innovative and unique approach and content, will meet the needs of all three audiences.

The overall plan for the book is both simple and logical. The first part of the opening chapter provides a thorough and updated overview of today's Electronics Industry: how the industry is organized, recent trends affecting job opportunities, and job qualifications and rewards for production workers, supervisors, equipment servicers, and field engineers. The second part of the same chapter focuses on the history of electronics from its earliest beginnings to the present, including the major contributions of great pioneers in the field and important breakthroughs such as the vacuum tube, the transistor, and the integrated circuit.

This introductory material is followed by a series of chapters dealing with electronics theory. These "concept chapters" provide a detailed and well-balanced background upon which a further knowledge of electronics theory may be built. A middle-ground approach is taken in these chapters so that the occupation-oriented readers may find real substance while avocationally oriented readers remain interested. That is, electronics theory and concepts have been stated clearly, simply, and logically so that even a novice in this field can readily grasp and understand them. At the same time, concepts have been carefully related to practical examples so that future electronics technicians gain substantial background information on electromagnetism, DC and AC theory, and basic instrumentation and storage devices.

The most compelling chapters of the book deal with the wide variety of applications in electronics, including electronics in medicine and health occupations, in space and world exploration, in telecommunications, in recording and audio systems, in video systems and music, and in computers and other electronic devices used in the home, business, and industry. These "applications chapters" reinforce the theme of continuing change in both the technologies and the products described.

For example, in the field of medicine, the recent development of the nuclear magnetic resonance scanner is considered to be the greatest breakthrough in diagnostic medicine in the last century. Mean-

while, NASA's whole approach to space exploration has been changed by its successful series of reusable space shuttle vehicles—an accomplishment that now makes practical a manned space station. In the field of telecommunications, cable television, providing a variety of programs from distant production centers, is aggressively competing with standard television stations for its share of the viewing audience. Geosynchronous communications satellites that link these cable systems to their program sources also bring a variety of programming to private homes that are equipped with receiving dishes to intercept the microwave transmissions. Video cassette recorders are selling much more briskly than originally predicted, especially since Japanese imports forced sharp reductions in retail prices. Video discs are also selling well, having been handicapped at first by high equipment costs and limited program offerings. In the field of audio, the most exciting development has been the digital audio disc with its great dynamic range and extended frequency response. Read by a laser beam and sealed in plastic, the disc should provide flawless reproduction for many years. In the home, the best-selling electronic equipment items have been video games and personal computers. In business and industry, two important electronic technologies are emerging. The first is optoelectronics—the technology of data transmission over a modulated beam of light and then over an optical fiber to its terminal. The second is robotics—the use of a microprocessor-controlled arm to perform simple, repetitive tasks such as assembling modules or welding and spray-painting automobiles. Thus, the second edition of *Electronics: Concepts, Applications, and History* correctly reflects electronics as it is today—exciting and expanding.

Particular mention should be made of the high level of public participation and active support received in the preparation of the applications chapters. One of the largest regional hospitals in the Pacific Northwest arranged for its own staff photographer to take many new photos to illustrate the chapter on electronics in medicine. At the same time, local medical doctors provided charts, pictures, and data in their specialty fields.

Similarly, a 50 kilowatt clear-channel broadcast station and a local full-power commercial television station assigned staff members to plan and then take the many black-and-white photos needed to illustrate the chapter on electronics in telecommunications. And finally, such prestigious national firms as General Electric, Heathkit, Hewlett-Packard, IBM, RCA, and Xerox, to name a few, searched their files for both printed materials and photographs that would be helpful in the writing of the text for other applications chapters and providing suitable illustrations.

The author also wishes to acknowledge the full support of Sacred Heart General Hospital of Eugene, OR, which assigned Jay Rymeski, their director of audio-visual instruction, to the taking of several dozen photographs needed to illustrate the chapter on medical electronics. Also appreciated is the assistance of Dr. John N. Mundall, who provided examples of normal electroencephalograms, and that of Dr. Stanley M. Richmond, also of Eugene, who supplied information on the technology of acupuncture. Because of the highly specialized nature of the subject matter of Chapter 8, the author wishes to acknowledge the contributions of the following people: Ruth McCarroll, Radiologist-Technician at Eugene Hospital and Clinic, Eugene, Oregon; Tom Lawry, Director of Community Relations and the Medical-Electronics Staff at Sacred Heart General Hospital, Eugene, Oregon; and David Kelch of the Medical Products Division of Hewlett-Packard—all of whom read the original chapter and supplied suggestions and photos of new equipment. Special thanks to Dr. Daniel Robinhold, Cardiologist, for supplying several types of heart pacer to Alan Yordy, Media Coordinator of Sacred Heart General Hospital, who took the photos of these units that appear in the text.

ELECTRONICS:

Concepts, Applications, and History

Contents

PREFACE xiii

1 ELECTRONICS TODAY AND YESTERDAY 1

The Field in Perspective 2
Today's Electronics Industry 2
How the Industry Is Organized 3
Recent Trends Affecting Job Opportunities 3
Two Emerging Technologies 5
Plant Organization 6
Job Opportunities and Incentives 7
History of Electronics 7
Early Developments 8
Pioneers in Electricity and Electronics 8
Summary 17
Key Words and Concepts 18
Self-Test Questions 18

2 BASIC CONCEPTS AND DEFINITIONS 19

Atomic Theory of Matter 20
Elements 20
Mass of Atoms 21
Free Electrons 21
Ions and the Generation of Electricity 22
Fission 23
Fusion 23
Compounds 24
The Concept of an Electrical Circuit 24
Current Flow in the Basic Circuit 25
Electrical Power 26
Ways of Generating Electrical Pressure 27
Summary 29
Key Words and Concepts 30
Self-Test Questions 30

3 BASIC LAWS, CIRCUITS, AND DEVICES 31

Coulomb's Law of Charges 32
Kirchhoff's Laws 33
Ohm's Law 35
Three Basic Statements of Ohm's Law 35
Problems and Solutions 36
Power Computations 37
Power and Wattage 37
The Time Factor 37
Three Basic Formulas for Computing Power 38
Electrical Circuits 38
Open and Short Circuits 38
Series Circuits 39
Parallel Circuits 40
Series-Parallel Circuits 41
Resistance and Resistors 42
Resistive Materials 42
Carbon Resistors 43
Wire-Wound Resistors 43
Variable Resistors 43
Color Coding Resistors 45
Summary 46
Key Words and Concepts 47
Self-Test Questions 47

4 BASIC ELECTROMAGNETISM AND DC THEORY 49

Magnetism 50
- Magnetic Poles 50
- Magnetic Materials 51
- Temporary and Permanent Magnets 51
- Magnetic Field Behavior 52
- Theories of Magnetism 53
- Magnets as Complete Systems 53
- Types of Permanent Magnets 54

Electromagnetism 55
- Strength of an Electromagnet 55
- Applications of Electromagnetism 56

Inductance 56
- How Inductance Works 56
- Inductive Reactance 58
- Self-Inductance and Mutual Inductance 59
- Solenoids 59
- Measuring Inductance 61

Capacitance 61
- Factors That Determine Capacitance 61
- Variable Capacitors 62
- Charging and Discharging Capacitors 62
- Units of Capacitance 63
- Naming Capacitors 63
- Electrolytic Capacitors 64
- Voltage Rating of Capacitors 64
- Capacitors in Series and in Parallel 64
- Precautions in Connecting Capacitors 65
- Capacitive Reactance 65

Combining Inductors and Capacitors 67
- Tuning and Tuned Circuits 67
- Calculating Inductance and Capacitance 68

Summary 68
Key Words and Concepts 69
Self-Test Questions 69

5 BASIC AC THEORY AND DEVICES 70

Some Applications of Alternating Current 71
Overview of AC 72
- How DC Differs from AC 72
- AC as a Power Source 74
- How AC Flow Is Produced 74
- Commercial Generators and Alternators 74
- AC Voltage and Current Waveforms 76
- Improving Alternator Efficiency 78
- Peak Values and Effective Values 79
- Impedance 80

Electric Motors 81
Transformers 82
- Step-Up and Step-Down Ratios 82
- Structure of Power and Audio Transformers 83
- Applications of Transformers 85

Summary 86
Key Words and Concepts 87
Self-Test Questions 88

6 BASIC INSTRUMENTATION AND STORAGE DEVICES 89

The D'Arsonval Meter 90
- Construction of the D'Arsonval Meter 90
- Operation of the D'Arsonval Meter 92
- Measuring Electrical Pressure 93
- Using the Meter on AC 94
- Measuring Current Strength 95
- Measuring Heavy Currents (AC) 96
- Connecting Voltmeters and Ammeters 96
- Measuring Resistance and Continuity 97

Volt-Ohm-Milliammeters 98
- Reading a VOM Dial Scale 98
- FET Voltmeters 99
- Digital Voltmeters 101
- Reading Power in a Circuit 101
- Measuring Power at Radio Frequencies 102
- Oscilloscopes 103

Cells and Batteries 103
- Carbon-Zinc Cells 104
- Rechargeable Cells 105
- Miniaturized Cells 105
- Charging Batteries 106

Summary 107
Key Words and Concepts 108
Self-Test Questions 108

7 AN INTRODUCTION TO ELECTRONIC COMMUNICATIONS 109

Transducers 110
- Common Transducers 110
- Transducer Efficiency 111

Active Devices 112
- Edison's Contribution 112
- The Fleming Valve 113
- The Triode Vacuum Tube 114
- The Vacuum Tube as a Variable Resistance 115
- Newer Active Devices 116
- Comparison of Active Devices 117
- Active Devices as Variable Resistances 119
- Integrated Circuits 119
- Oscillation 122

Electronic Communications 125
- Historical Perspective 125
- The Electromagnetic Spectrum 126
- Propagation of Radio Waves 127
- Impact of Sunspots and Sunspot Cycles 129
- Conveying Information by Electronic Communication 130

Summary 132
Key Words and Concepts 134
Self-Test Questions 134

8 ELECTRONICS IN MEDICINE AND HEALTH OCCUPATIONS 135

Historical Perspective 136
- The Curie Family 136
- Wilhelm Roentgen 136
- Arsene D'Arsonval 137

Overview of Medical Electronics 139
- Operating Room Equipment 139
- Emergency Room Equipment 140
- Intensive Care Equipment 141
- Coronary Care Equipment 141
- General Patient Care Equipment 142

Electronic Equipment for Diagnosis 143
- X-Ray Machines 143
- CAT Scanners 146
- NMR Scanners 148
- Other Scanning Machines 150
- Ultrasound Equipment 151
- Electroencephalograph 152
- Electromyograph 153
- Electrocardiograph 154
- Thermographic Equipment 155
- Electron Microscope 156

Electronic Equipment for Therapy 157
- Diathermy 157
- Ultrasound 158
- Acupuncture 159

Electronic Life-Support Equipment 160
Maintaining Medical-Electronic Equipment 161
Summary 162
Key Words and Concepts 163
Self-Test Questions 163

9 ELECTRONICS IN SPACE AND WORLD EXPLORATION 165

Historical Perspective 166
An Overview of NASA Electronics 166
- Basic Requirements for Space-Vehicle Electronics 166
- Basic Requirements for Ground-Station Electronics 167
- The Space Program's Effect on Electronics 169

Highlights of NASA's Space Program 170
- Early Period of Space Exploration 170
- Space Probes 170
- Space Shuttle Orbitors 177

NOAA 178
- NESDIS 179
- Weather and Environmental Satellites 179
- Landsat and the Multispectral Scanner 181

Summary 182
Key Words and Concepts 182
Self-Test Questions 183

10 ELECTRONICS IN TELECOMMUNICATIONS 184

Historical Perspective 185
AM Broadcast Stations 186
Classification of AM Stations 186
Advantages of AM Transmission 187
Job Assignments in AM Stations 187
FM Broadcast Stations 188
Development of FM Transmission 188
Growth of FM Broadcasting 190
Classification of FM Stations 190
Television Broadcasting 190
Early Television Technology 190
How Television Works 191
Emergence of TV Standards 192
Color Television 193
Television Job Assignments 194
Cable Television 194
Use of Geosynchronous Satellites 195
Cable TV Installation 197
Rooftop Reception of Satellite TV 197
Future of Radio and TV Broadcasting 199
Future of AM Stereo 199
Future of Commercial TV Stations 199
Amateur Radio 200
Contributions of the Amateur Radio Service 200
Amateur Radio Today 200
OSCAR Space Satellites 201
American Radio Relay League 201
Citizens Radio Service 202
Problems Facing CB Radio 202
Radio Emergency Associated Citizens Teams 203
CB as a Training Ground for Amateur Operators 203
Summary 203
Key Words and Concepts 204
Self-Test Questions 204

11 RECORDING AND AUDIO SYSTEMS 206

Recording and Reproduction of Sound 207
Edison Phonograph 208
Edison's Recording Unit 208
Record Discs 208
Sound-Motion Pictures 209
Electronics and the Recording Process 209
Union of Radio and Phonograph 210
Improvements in Record Discs 210
Record Disc Speeds 210
Magnetic Recording and Playback 211
Earliest Tape Recorders 211
How Magnetic Tape Recording Works 212
Improvements in the Basic Unit 212
Audio System Components 213
Amplifiers 213
Loudspeakers 215
Pickup Cartridges and Styli 216
Microphones 217
AM/FM Tuners 217
Other Accessory Equipment 218
Room Acoustics 218
Developments in Recording and Playback 219
New Coatings for Tapes 219
Cassettes 219
Digital Technology 219
Summary 221
Key Words and Concepts 222
Self-Test Questions 222

12 VIDEO SYSTEMS AND ELECTRONICS IN MUSIC 223

Video Tape Recording and Playback 224
Review of Magnetic Recording 224
Video Recording and Frequency Response 225
Helical Scan Tape Recorders 226
Video Disc Recording and Playback 229
A Review of Lasers 229
Recording and Playback Process 229
Playback Alternatives 230
Electronics in Music 231
First Electronic Musical Instrument 232
Electromagnetic Pickups 232
Electronic Organs 233
Electronic Pianos 234
Music Synthesizers 235
Summary 235

Key Words and Concepts 237
Self-Test Questions 237

13 ELECTRONIC COMPUTERS, CALCULATORS, AND CLOCKS 238

Historical Perspective on Computers 239
European Advances 239
Babbage's Analytic Engine 239
Development during World War II 240
Introduction of Transistors 240
Microcomputers 240
Categories of Computers 241
Digital Computers 241
Basics of Computer Design 241
Building Blocks of the Computer 242
Home Computer Installations 247
Microprocessors 248
Electronic Calculators 249
Construction of Electronic Calculators 249
Trends in Visual Displays 250
Trends in Calculators 250
Clocks and Watches 250
Concept of Timekeeping 251
Construction of Electronic Clocks 251
Electronic Watches 253
Summary 253
Key Words and Concepts 254
Self-Test Questions 255

14 GENERATION AND DISTRIBUTION OF ELECTRICAL POWER 256

Historical Perspective 257
First Power Station 258
Introduction of AC Power Generation 258
Basic Concepts of Power Generation 258
Power Generation 258
Transmission Lines 259
Typical Power Distribution System 261
Steam Plants 262
Problems of AC Power Distribution 262
Distribution of DC Power 264
Principles of DC Transmission 264
Converting DC Back to AC 265
Future Trends 266
Nuclear Power Generation 266
Atomic Fusion 266
Atomic Fission 267
Nuclear Power Plants 267
Alternate Sources of Electrical Energy 268
Solar Energy 268
Wind Power 269
Metglas and AC Power Distribution 271
Summary 271
Key Words and Concepts 272
Self-Test Questions 272

15 ELECTRONICS TODAY IN THE HOME, BUSINESS, AND INDUSTRY 274

Electronics in the Home 275
Microwave Ovens 275
Solid-State Laundry Units 276
Smoke Detectors 277
Photography and Cameras 277
Electronic Weather Stations 279
Color Projection Television 280
Video Games 281
Cordless Telephones 282
Electronics in Business and Industry 284
Telephone Answering Systems 284
Electrostatic Copying 284
Word Processing 285
Robots and Robotics 287
Optoelectronics 288
Latest Trends in Electronics 289
Summary 290
Key Words and Concepts 291
Self-Test Questions 291

16 A LOOK AT THE FUTURE OF ELECTRONICS 292

Five-Year Forecast for Electronics 293
Industrial Production 293
Industrial Employment 293

Five-Year Forecast for New Applications 295
Medicine and the Health Occupations 295
Space and World Exploration 296
Telecommunications 296
Audio Systems 297
Video Systems 298
Electronic Musical Instruments 298
Computers 299
Calculators 299
Electronic Clocks and Watches 300
Power Generation and Distribution 300
New Uses for Electronics in the Home 301
New Uses for Electronics in Business and Industry 302
Automotive Electronics 303

INDEX 305

1 Electronics Today and Yesterday

This woman is working in electronics drafting. Women readily find employment with electronics firms in many job assignments. (Photo courtesy of Tektronix)

The Field in Perspective

Today's Electronics Industry
- How the Industry Is Organized
- Recent Trends Affecting Job Opportunities
- Two Emerging Technologies
- Plant Organization
- Job Opportunities and Incentives

History of Electronics
- Early Developments
- Pioneers in Electricity and Electronics

Summary

Key Words and Concepts

Self-Test Questions

The purpose of this chapter is to provide some insight into the importance and scope of electronics in terms of both its economic impact and its far-reaching effects on our lives. Electronics is an industry of great vigor and promise, and each step forward in its technology will lead to further important applications. These applications, in turn, will result in new jobs in the design, construction, and sale of new products.

Perhaps the most remarkable characteristic of the Electronics Industry is its tremendous expansion. Since 1920, annual gross income has risen from just a few million dollars for a handful of products to many billions of dollars for an enormous variety of products and services. Moreover, this growth has taken place over a period of less than sixty years—a record of expansion that other fields of industry have seldom equaled.

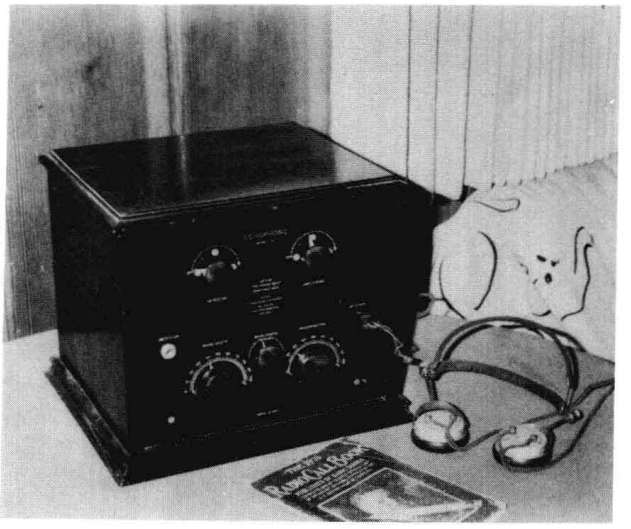


Figure 1.1 A Mid-1920s Home "Listening Post" It consisted of a battery-powered broadcast receiver connected to headphones. The 1926 *Radio Callbook* was necessary because stations operated on irregular schedules and often changed frequencies.

THE FIELD IN PERSPECTIVE

Many people can recall the almost complete lack of electronics and its applications in American life in 1920. For example, the longest distance a telephone call could travel was from New York to Denver, and connections were uncertain. What little radio broadcasting took place did so with low power and on an erratic schedule. The focal point for home entertainment was the windup acoustic phonograph with its cumbersome shellac records and rapidly wearing steel stylus. Motion pictures were of the silent type, with the dialogue projected on the screen between scenes. Hospitals had little or no electronic equipment for diagnosis, therapy, or monitoring.

The sudden upsurge of radio broadcasting in 1922 sparked rapid development in the Electronics Industry. That year marked the first large-scale use of the three-element vacuum tube (triode) invented earlier by Lee De Forest. During the period from 1920 to 1923, the number of broadcast stations leaped from 3 to more than 500, and sales of radio receivers (Figure 1.1) skyrocketed from $2 million to $136 million. General Electric, RCA, and Westinghouse, anxious to keep their receiver production lines operating at full capacity, constructed, operated, and programmed their own broadcast stations. Their top-quality programs stimulated additional set purchases. With a ready supply of mass-produced radio components, including vacuum tubes, resourceful technicians and experimenters were soon developing innovative circuitry for entirely new applications. Electronics was off to a whirlwind start, and the pace of development has never slackened.

TODAY'S ELECTRONICS INDUSTRY

Characteristic of today's Electronics Industry are steady and rapid growth; widespread automation, particularly in the production of integrated circuits; high worker productivity; and a comparatively small but highly trained work force. How big is the in-

dustry in terms of dollar volume and in terms of the number of persons employed? The *U.S. Industrial Outlook for 1982*, published by the Department of Commerce, estimated the total value of industry shipments by American manufacturers of Electronic Equipment and Electronic Components (they are considered separate divisions) at $42.4 billion—an increase of 8.9% from the 1981 figure of $38.9 billion. Much of this increase can be attributed to greatly expanded purchasing by the Department of Defense, which uses an enormous spectrum of electronic systems.

Current information from the Department of Commerce indicates that the Electronics Industry is relatively depression free. For example, during the downturn years of 1981–1982, workers in many major industries received wage cuts or even termination notices; in electronics, however, the number of workers actually expanded slightly, and many of them received wage increases averaging slightly under 10% per year.

In preparing for employment in electronics, you need to keep in mind the importance of choosing a rapidly growing segment of the industry. Choosing an area that is static or is in decline may result in a dead-end job without a future. The following sections examine the more promising areas of electronics employment.

How the Industry Is Organized

The Department of Commerce divides the Electronics Industry into two major categories and, at the same time, identifies perhaps half a dozen smaller divisions. The first major division is Electronic Equipment, which comprises a large number of electronic assemblies produced from such components as transistors, resistors, and capacitors. The second major division, Electronic Components, represents the manufacturing plants that produce the parts used in equipment fabrication.

The category of Electronic Equipment contains two subdivisions that indicate the eventual user of the product. Electronic Systems (Government Equipment), the larger of the two divisions, includes purchases made by the Department of Defense, NASA, NOAA, FAA, and other government agencies. Consumer Electronics includes equipment destined for use in the home or by individual consumers—for example, microwave ovens, video cassette recorders, TV receivers, and personal computers.

Between 1978 and 1983, sales of electronic equipment have grown from about $14 billion to $18 billion. During these years, the number of skilled workers in the field has increased from 300,000 to 325,000, and each year the field has shown a substantial increase in the value of its shipments. Over the same five-year period, sales of electronic components have increased from $10 billion to $27 billion, and employment figures have increased from 309,000 to 504,000. Also, shipments of repair parts to workers who service electronic equipment have been steadily increasing.

Recent Trends Affecting Job Opportunities

Certain developments that began in mid-1983 and carried over into 1984 may have a far-reaching effect on Electronics Industry employment. Late in the spring of 1983, video games and video game machines, which had been selling so briskly that it was difficult for retailers to keep them in stock, suddenly encountered great sales resistance. Prices tumbled overnight, but sales remained abnormally low. Major firms such as Atari and Mattel, which had specialized in video games, suffered severe financial losses. Atari sought to reduce production costs by closing a number of its U.S. plants and moving manufacture of its game machines to the Far East—a decision that cost hundreds of American workers their jobs. Were other American manufacturers to follow the same course, the effects on employment in the Electronics Industry would be very unsettling.

In the early summer of 1983, a similarly abrupt drop in personal computer sales caught retailers in a bind worsened because factory shipments of com-

puters were increasing each month. Within a few weeks retail prices of well-known PCs tumbled—in some cases to less than half their previous level. A number of smaller computer firms declared bankruptcy. Others cut back their work force drastically, and thousands of skilled workers lost their jobs. Today the demand for PCs can be described only as fair, despite the bargains available at computer stores. Manufacturers claim they are losing money on every sale, and some companies may eventually go bankrupt. These developments have made it inadvisable to look for employment in the video games or personal computer segments of the Electronics Industry.

Job Qualifications for Production Workers

Production workers account for approximately 80% of an electronic assembly plant's employees. They are usually described as hi-tech workers. Most employers expect potential employees to have a high school diploma or its equivalent, but few applicants will be turned down if they seem bright and eager. Previous experience in electronics is usually not required, although some previous training in the field may be helpful when the number of job applicants exceeds the number of vacancies.

In a typical assembly line, workers sit at a bench equipped with hand tools and bins of small electronic parts. A completed model of the circuit board on which they are working is prominently displayed in front of them. The worker must bend leads on small parts and then insert the leads into predrilled holes in the circuit board. After clipping off the excess lead length, the worker solders the component in place on the circuit board. An excellent job of soldering is required. Besides soldering and hand-tool skills, workers need only be familiar with the resistor and diode color codes.

Job Qualifications for Supervisors

The educational background and the skills needed by upper-echelon workers in electronic systems or electronic components plants are different from the requirements established for production workers. In general, production-line supervisors and department heads need a minimum of two years of electronics training, both theory and shop skills, and some successful supervisory experience. Chief engineers and their assistants should hold a degree either in electrical or in electronic engineering from an accredited college. They must also be able to design sophisticated electronic equipment.

Job Qualifications for Equipment Servicers

The armed forces train selected personnel for electronic equipment repair, but civilian repairers are far more numerous. Personnel engaged in electronic equipment repair number about four or five times

Hewlett-Packard engineers ponder changes in a circuit board and its supporting structure, an essential module of the professional oscilloscopes shown in the foreground. Engineers are always alert to changes that will improve equipment performance and reliability. (Photo courtesy of Hewlett-Packard)

the 325,000 workers employed in electronic equipment construction. However, the number of repairers is decreasing for two reasons. First, increased technology, particularly the trend toward integrated circuits, is forcing out the part-time tinkerers who were largely hobby oriented and had little or no formal training. Second, the changing economy of the marketplace has made the traditional "one-person shop" no longer practical. A large capital outlay is required for the purchase of adequate test equipment, and a large inventory of various integrated circuits is needed.

Shop technicians must attend a vocational school or a community college for at least one year of formal training in electronics theory and shop practices; two years of training is preferable. Shop skills that should be stressed include signal tracing; interpretation of oscilloscope patterns; and the use of modern pieces of test equipment such as radio frequency and audio frequency signal generators, distortion analyzers, capacitor testers, and volt-ohmmeters. Courses in physics and math are also recommended, since electronics is frequently described as a highly specialized field of physics. Avocational and vocational experience is frequently helpful too. For example, technicians who have worked as amateur radio operators or as chief engineers or assistants in radio or television stations often find their background helpful in equipment servicing.

Newer items of consumer electronics—for example, video cassette recorders and digital disc players—frequently present repairers with mechanical as well as with electronic problems. Therefore, skill in using small hand tools is frequently helpful, as is the ability to spot a mechanical problem quickly. (Such ability is perhaps an inborn talent; it is almost impossible to teach.)

Prospects for employment in electronic equipment servicing are mixed. Retail outlets that deal in consumer electronics need well-staffed service departments. Also, as more complex technology has forced less competent people out of the servicing field, wages have risen accordingly, and prospects for future wage increases are good. However, as wages rise, repair becomes less practical. Many electronic items become throwaways because it is cheaper in the long run to buy a new unit than to repair the old one.

Job Qualifications for Field Engineers

Firms that sell both personal and smaller business computers have been responsible for the emergence of a new job category: the field engineer. Because installation of computers involves peripheral equipment such as video monitors, keyboards, disc readers, and printers, servicing at a central location would be too difficult. Therefore, when equipment malfunctions, a field engineer calls on the computer's owner and analyzes the problem. The field engineer then reports the nature of the difficulty to the computer firm's service manager, who dispatches a repair person. Field engineers must be articulate and presentable and should have at least a year of technical training at a community college or vocational school. The field engineer concept has increased public acceptance of new product lines that have not achieved name recognition.

Two Emerging Technologies

The Department of Commerce predicts that two emerging technologies—optoelectronics and satellite communications—will show rapid growth and development over the next five years. *Optoelectronics* involves transmission of many types of data over optical fibers of quartz, glass, or plastic by means of a beam of light on which information has been impressed, or modulated, at the transmitter unit. At the far end of the optical fiber, a receiver unit recovers the data from the varying light beam. Telephone companies are experimenting with a similar technology to transmit conversations over an extended distance. As a light source they use a modulated **laser beam**—a device that transforms light of varying wavelengths into a narrow and intense beam of *coherent light* (light of a single frequency).

Satellite communications refers to the use of *geosynchronous satellites*—that is, satellites synchronized with the earth's rotational speed. They

orbit nearly 23,000 miles above the equator and pick up data-bearing radio signals aimed at them by ground stations. The satellites then retransmit the signals to suitably equipped ground stations dispersed over a large geographical area. Such satellites already bring us network radio and TV programs and cable TV programs from program production centers. In the future they will bring computer networking and a computerized news and shopping service into many homes.

Plant Organization

Research and Development

In an electronics manufacturing plant, the department of research and development (R&D) determines new product needs and develops new products to the point of engineering practicality. R&D then gives the model, usually called a prototype, to the production department for large-scale manufacture. R&D is always headed by a chief engineer with a degree in electrical engineering or some related area, plus extensive practical experience in electronics. Chief engineers design equipment that will work effectively and reliably and can be produced and sold in a highly competitive market.

Assisting a chief engineer are one or more electronics technicians, who take a design and from it develop the first model, or prototype. The technicians then work on perfecting the model until it's ready for mass production. Technicians' training includes a degree in electronics from a community college or a vocational school, plus several years of experience in some phase of electronics. Additional classes and increased experience may lead to promotion to engineering positions.

Production

The step after product design is production, or manufacturing, which involves ordering and testing of components from other manufacturers, assembling of components, and final testing and calibration. The production department is usually headed by a person with extensive supervisory experience in the manufacture of electronics or in related fields. A typical production line includes about as many women as men. Most workers hold at least an associate of science degree from a community college or a vocational school and receive advanced training in electronics and technology in classes conducted by their employer.

After equipment reaches the end of the production line, it receives careful inspection, testing, and final checkout and calibration by a team of junior engineers and advanced electronics technicians. An occasional sample is taken from the production line and sent to quality control for a detailed inspection that reveals quality of workmanship and resistance to mechanical vibration and to extremes of temperature and humidity. Production departments try to keep the ratio of engineering to technician personnel in both production and quality control areas to no more than 1:3.

Customer Service

Also known as field service or field engineering, the purpose of the customer service department is to keep customers happy. Members of this department are basically sales personnel with some intensive training in electronics, particularly as it applies to the employer's product. Traveling by air, train, or car, field personnel call on firms that have purchased the company's product. Customer service representatives check on whether minor adjustments can be made to improve products. They also make their employers aware of new designs that could improve their company's product or service.

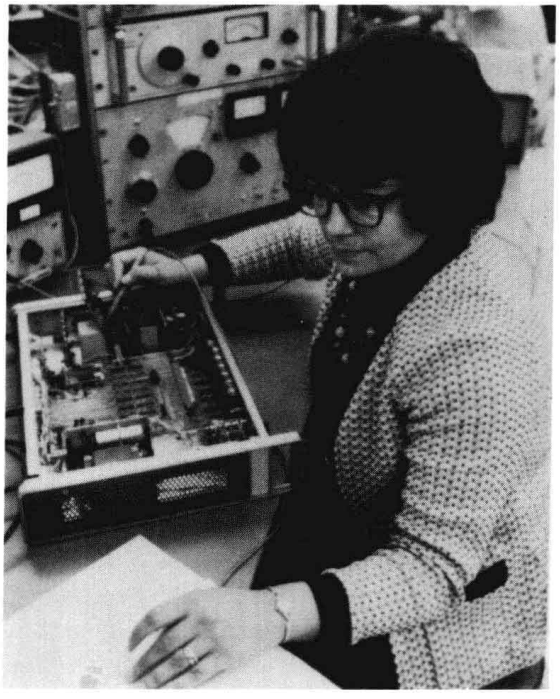
Supportive Services

Any manufacturing firm must also have supportive services such as purchasing, accounting, sales, and public relations. All these departments are coordinated by management. Of particular importance in electronics is the educational department, which trains new employees to turn out acceptable work and encourages other staff members to improve their skills.

Job Opportunities and Incentives

A typical electronics firm today offers a wide spectrum of jobs, both in technological areas and in more traditional assignments. These firms also show a consistent annual growth rate of 7% or better. They offer attractive wage scales and fringe benefits because of brisk competition for qualified personnel at the entry level and because of the advantage that accrues to a firm that retains highly skilled staff members dedicated to their assignments.

In November 1983, major West Coast electronics firms were paying beginning workers with minimal background at least $6.00 an hour. The average wage of a trained worker was often more than twice this amount. Senior engineers with a doctoral degree and years of successful designing experience were earning upwards of $40,000 a year. Fringe benefits included free or low-cost training classes leading to both job promotions and advanced degrees, prepaid medical and dental insurance, recreational programs, and a retirement plan in which the firm matched the employee's contributions. Employees of junior engineer rank or better were offered stock-purchase options, and some companies included all staff members in profit-sharing plans.



This woman is enrolled in an apprenticeship program at Hewlett-Packard. Such programs allow employees to upgrade job skills and qualify for promotions. (Photo courtesy of Hewlett-Packard)

Two large electronics firms on the West Coast encourage applicants by offering to waive rigid hours of employment and by granting attractive fringe benefits. Their approaches are somewhat different. Hewlett-Packard of Palo Alto, California, offers a flexible starting-hour schedule for a shift of regular length. Tektronix, near Portland, Oregon, has part-time workers report for work at regular plant hours but assigns them to the production of modules, or subassemblies. It is therefore practical for employees to work as few as three or four hours or as long as a normal shift. The principal fringe benefit offered by each firm is the opportunity for workers to attend special low-cost or free-tuition classes to upgrade their skills—an opportunity that can lead to full-time professional employment later on.

HISTORY OF ELECTRONICS

This section discusses the development of electronics, its pioneers and their principal contributions, and some of the breakthroughs that brought about rapid growth of the field.

Newcomers to the field of electronics frequently ask: "What is the difference between electricity and electronics?" The best answer is to give an acceptable definition of both. Considerable overlap exists between these two closely related fields, and persons employed in electronics are expected to have both theoretical and practical knowledge of electricity and its many applications.

Electricity: The movement of electrons through such elements as conductors, inductors, capacitors, and resistors to produce heat, light, and motion.

Electronics: The branch of electricity that deals with the behavior of electrons—particularly those of the substances we call semiconductors—in a vacuum, in gases, in liquids, and in crystals.

Early Developments

Our earliest knowledge of electricity and magnetism came from the Middle East and China, where both mathematics and astronomy were well developed long before their introduction to Europe. The compass was widely used in China by A.D. 21, although early compasses pointed toward the south pole rather than toward the north.

The ancient Greeks, too, left us a written record of their early contributions to our knowledge of magnets, as well as their first speculations about the composition of matter. As far back as 600 B.C., they had noted that, if first rubbed with a cloth, lumps of fossilized resin from ancient trees would attract dust, pieces of straw, and other materials. They called this resin *elektron*. They also noted that in nearby Crete and Asia Minor, where iron ore deposits lay on the surface, certain lumps of the material would attract objects made of iron. These same naturally occurring magnets, when suspended from a thread, would point toward geographic north and south. The ancient Greeks called these magnets *lodestones*, which means "leading stones." As far back as 460 B.C., Democritus suggested that all matter was formed of infinitely tiny particles that he called *atomonon*, the root of our present term *atom*.

Pioneers in Electricity and Electronics

What follows is more than a chronology of scientists and their contributions to our knowledge of electricity and electronics. It is an attempt to show how each individual's thinking and contributions affected the thinking and contributions of other scientists who were exploring the field at about the same time. Many of the pioneers are discussed in greater depth in the applications chapters. Edison, Fleming, De Forest, and Armstrong, for example, are not included in the following sequence.

Benjamin Franklin (1706–1790)

Benjamin Franklin was a remarkably talented American who seemed to be a success at every enterprise in which he became interested. Widely known as a writer, editor, and statesman, he was intrigued by all aspects of physical science, including the relatively new field of electricity.

Legendary is the famous experiment in which Franklin attached an iron key to the tail of a kite and sent it aloft during a violent thunderstorm. When he brought it down, he found the key magnetized—evidence that electricity and magnetism were closely related. Franklin also experimented with the static charges on amber and glass rods. He found that, after being rubbed with a cloth, amber rods repel amber rods and glass rods repel glass rods, while an amber and a glass rod attract each other. He concluded that the same physical laws that apply to magnets also apply to electric charges: Like charges repel, and unlike charges attract.

Franklin also concluded that an electric current resembles a fluid as it moves through the circuit from positive (a point where the fluid is in excess) to negative (the point of deficiency). This positive-to-negative concept of electrical current flow found wide acceptance among other experimenters. Even today apprentice electricians use the concept in their studies. In electronics, however, current flow is generally regarded as produced by a progressive drift

of electrons from the negative to the positive pole of the circuit, and this book takes that view.

Hans Christian Oersted (1777–1851)

Although Franklin had theorized about a direct relationship between electricity and magnetism, it remained for the Danish scientist Hans Christian Oersted to establish without question that electricity and magnetism are related. In a classroom demonstration, he showed that when a conventional magnetic compass was brought near a wire through which direct current was flowing, the compass needle was strongly deflected. This compass deflection continued until the current was turned off, at which time the compass was once more affected by the earth's magnetic field and pointed roughly north and south. When the current direction in the wire was reversed, the compass needle was deflected again, this time in the opposite direction.

Oersted, because of limited professional training and a lack of mechanical skills, was not capable of advanced research, but his published findings set off a flurry of activity among better-qualified experimenters such as Ampere, Faraday, and Henry.

Michael Faraday (1791–1867)

Because he was the son of an English blacksmith, a formal education was out of the question for Michael Faraday, and he was apprenticed to a London bookbinder. Fortunately, he had an employer who encouraged him to read some of the books he bound and also to attend the public lectures of the scientists who had written them. Through attendance at these lectures, he gained access to people engaged in laboratory research, and even to their laboratories.

One of Faraday's "firsts" was the development of the electric motor. A little later he produced the first iron-core power transformer, and in 1831, after several attempts, he developed the first **alternator,** or **generator**—a machine for producing electricity by magnetism. Invention of the alternator is often considered the greatest single electrical discovery of all time. All three of Faraday's major developments were stimulated by his reading about Oersted's experiments with electromagnetism.

George Ohm (1787–1854)

A science teacher in German high schools for most of his life, George Ohm desperately wanted to produce some spectacular research that would lead to a teaching appointment at a university. He was the first scientist to note the importance of **resistance** (the opposition a substance offers to the flow of an electrical current). His published findings, known as *Ohm's Law* (see Chapter 3), have proved to be the most important law of electricity and electronics.

Scientists in Great Britain nominated him as a member of the Royal Society, but in his own Germany scientists questioned his methods and findings, and it was seven years before he received a teaching appointment at the University of Munich. For his contribution, the unit of resistance is named the **ohm.**

Joseph Henry (1798–1878)

By the mid-19th century many of the great discoveries of the early period of electricity and electronics had already been made. Volta had invented the cell and the battery; Oersted had pointed out the relationships between electrical current flow and electromagnetic fields; Faraday had patented the motor, the transformer, and the generator; and Ohm had pointed out the relationships between electrical pressure, resistance, and current flow. Most later discoveries such as the telegraph, the telephone, and the radio, were to come in specific fields. Needed at this time was a person who could tie together the knowledge obtained from previous discoveries and encourage the scientific community to use the knowledge constructively in special applications. That man proved to be the American scientist Joseph Henry.

Like Faraday, Henry was impressed with Oersted's experiment relating electrical current flow to electromagnetic fields. One of his first projects was to analyze the factors determining the strength of a particular electromagnet. He found three basic

factors to consider: the number of turns of wire in the winding, the nature of the core material, and the strength of the electrical current flowing through the windings. His attention then shifted to the development of a basic telegraph system, which used an electromagnet at a distant point to attract an iron clapper that rang a bell. He also invented a device still widely used, called the *electrical relay,* an electromagnetically operated switch with a trickle of current through the windings controlling a much heavier current flowing through its contacts.

By 1831 Henry had developed all the elements of a practical telegraph system but he decided not to file a formal patent application for it. In the same year he analyzed Faraday's demonstration motor and decided he could greatly improve on it. The series-wound motor Henry developed is, in modified form, widely used today.

Many scientists recognized Henry's unique abilities both as a researcher and as an organizer, and upon their recommendations he was appointed executive secretary of the Smithsonian Institution. In this capacity he could counsel experimenters regarding their inventions and make small cash grants to experimenters in financial need. His office became a clearing house of information for the scientific world. Among the inventions he assisted in developing and promoting were Morse's telegraph and Bell's telephone.

Samuel Morse (1791–1872)

Samuel Morse had a most unlikely background for the developer of a communications system based on electromagnetism. His system broke with traditional emphasis on visual communication devices such as heliographs, semaphores, and signal flags. Although Morse had studied some physical science at Yale, he was trained as a portrait painter. However, he found his income inadequate for the needs of his rapidly growing family.

Morse's interest in communications was stirred by an acquaintance's remark that an impulse of electricity might travel for many miles along a metallic wire and could be used for signaling at a distance.

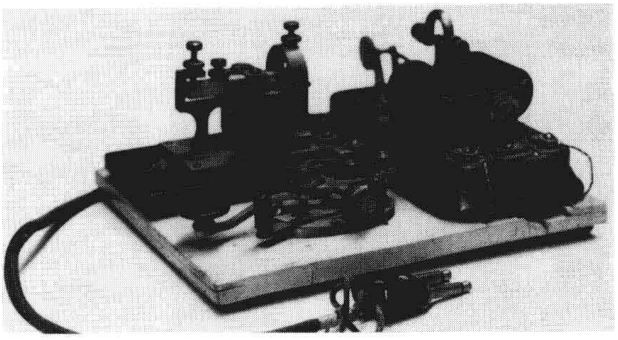


Figure 1.2 Morse Telegraph Equipment Used by Railways in the 1920s Separate transmitting and receiving units were provided.

He sought the advice of Joseph Henry, who generously shared with Morse information about the telegraph system he had developed. (Morse later denied that any such assistance had been given.)

Morse's principal contributions to the telegraph system were the system of dots and dashes known as the *Morse Code* and a piece of equipment that recorded incoming dots and dashes on a moving paper tape for later translation (Figure 1.2). Construction of the technical equipment, such as sending keys, electromagnets, and relays, he left to others with more technical skill.

Morse had a real flair for public relations, as shown by his demonstration of his telegraph system, financed by a grant from Congress, in 1844. His successful sending of a message from Baltimore to Washington, D.C., won him wide public acclaim. By the end of the Civil War, more than 200,000 miles of telegraph lines were carrying business and personal messages as well as regional and national news.

Alexander Graham Bell (1847–1922)

Alexander Graham Bell was a member of a Scottish family that had immigrated to America. Three generations of the family had worked with people who were hard-of-hearing or who had speech problems. His first experimentation was largely moti-

vated by the need for a device that would enable people who had partial deafness to hear adequately.

Unlike Morse, Bell was a skilled and trained experimenter and had a well-equipped and well-staffed laboratory at his disposal. His key staff member was Thomas Watson, who was particularly adept at quickly turning out well-designed equipment from Bell's hurried drawings and sketches.

Bell's basic premise for the invention of the telephone (the word means "sound at a distance") was as follows: If sound vibrations could be turned into a fluctuating electric current, the fluctuating current could be converted back to sound waves identical to the original vibrations.

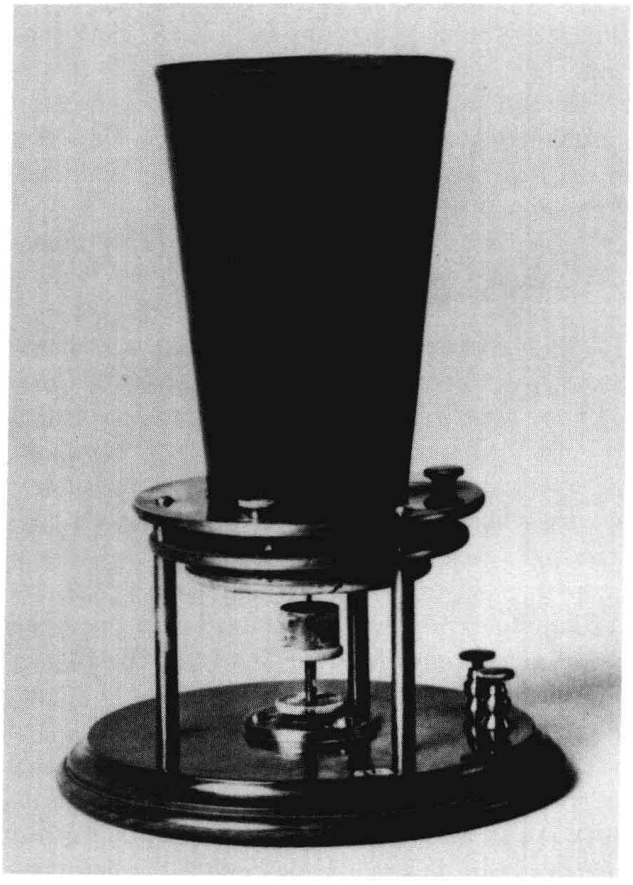


Figure 1.3 First Telephone Transmitter Bell's liquid microphone was in operation about 1876. The receiver is a magnetically vibrated metal reed. (Photo courtesy of Western Electric)

Once this premise had been clarified by discussions with Joseph Henry, Watson constructed two identical telephone units (Figure 1.3). Each unit had a light iron disc glued to a membrane made of sheep's intestines. This membrane was supported in a small wooden box immediately in front of an electromagnet (a steel magnet over which turns of wire had been wound). When Bell spoke into one unit, the membrane vibrated and moved the iron disc in the magnetic field, setting up a pulsating current in the turns of wire wrapped over the magnet. When these turns of wire were connected to the wire leads of the second unit, its vibrating membrane set up sound waves that, though faint, produced understandable speech.

Bell immediately applied for a patent, which was granted on 7 March 1876. It proved to be the most valuable patent in the history of the Patent Office. Later Thomas Edison and Emile Berliner were to invent the present-day carbon microphone, and all later models of the telephone have included their invention, which results in increased volume and better clarity of speech. Bell's next step was to separate the talking from the listening function to make the instrument easier to use. The result was a product similar to the present-day telephone.

Guglielmo Marconi (1874–1937)

Unlike many scientists and inventors, Guglielmo Marconi was comparatively wealthy. As a youth he was privately tutored and did not attend the public schools and universities of his native Italy. A hard-working experimenter who later became a shrewd businessman, he made three basic contributions to electronics.

First, he greatly improved the stability and power output of the spark transmitter invented by the German physicist and teacher Heinrich Hertz. He also replaced Hertz's air-gap receiver with a more effective device that used a coherer, a form of diode, as a detector. Second, he recognized that the trans-

mitting and receiving range could be enormously extended by using a long, elevated wire as an antenna and by using a low-resistance ground connection obtained by driving ground rods into the earth or by burying a series of wires just under the surface of the soil.

Marconi's third major contribution was his recognition of the commercial possibilities of electronics in both manufacturing and communications. At an early age, he moved to England and established the British Marconi Company, which engaged in sending wireless messages for a fee. It was also the first firm to manufacture electronic components and assemble them into transmitters and receivers for commercial sale.

After many months of preparation, including the installation of a powerful transmitter at Cornwall, England, and an elaborate receiving antenna at St. Johns, Newfoundland (Figure 1.4), Marconi was ready for a dramatic test of electronic communication. He transmitted the first radio signal across the Atlantic on 12 December 1901. After several unsuccessful attempts, the receiver operator at St. Johns heard clearly the three "dits" of the letter *s*, the agreed-on test signal. The Atlantic had been bridged, and Marconi soon established powerful wireless stations in New England and engaged in transatlantic wireless telegraphy.

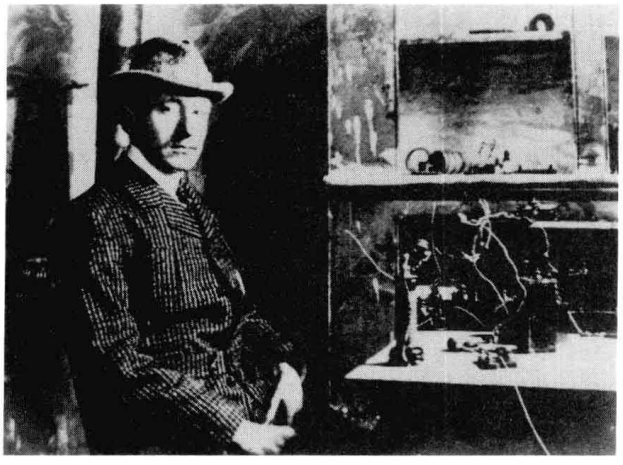


Figure 1.4 Guglielmo Marconi at His Radio Receiving Set at St. Johns, Newfoundland Here, he received the first transatlantic radio message in December 1901. (Photo courtesy of RCA)

Marconi also established the American Marconi Company in the United States to engage in the manufacture of radio receiving and transmitting equipment and to send and receive wireless messages for a fee. In 1919 the General Electric Company devised a plan by which the stock of American Marconi, already dominant in the U.S. marketplace, would be taken over completely by a new company to be known as the Radio Corporation of America, or RCA. The new firm purchased all the patents and assets of the American Marconi Company and entered into licensing agreements with General Electric, Westinghouse, and Western Electric. The takeover ended the threat of Marconi dominance in the United States in the fields of electronic manufacturing and electronic communications.

Nikola Tesla (1856–1943)

Nikola Tesla began his career as a scientist in Hungary and migrated to the United States in 1884, after which he made many discoveries in electricity and electronics that were to have a great impact on the American economy. After a brief period of employment with Thomas Edison, with whom he constantly disagreed, Tesla decided to concentrate on improving existing systems of alternating current generation, distribution, and utilization. Soon he had designed and constructed the first AC induction motor, which was based on the concept of rotating electromagnetic fields and was so effective that it became the prototype of most electric motors used to power both household appliances and large industrial machines. To provide an effective power source for his new motor, he devised the first practical polyphase (many currents flowing at the same time) system of AC power generation and distribution.

For a complete system, Tesla found it necessary to design and construct from scratch AC generators (alternators), step-up transformers, transmission lines,

step-down transformers, and many related items. George Westinghouse, an inventor who had made a fortune with his railroad air brake, gave financial help to Tesla and started to manufacture much of Tesla's equipment in his New York factory. Today, all major electrical utilities use Tesla's system of power generation and distribution.

Fascinated by Heinrich Hertz's experiments with the propagation and reception of radio waves through space without the use of wires, Tesla next focused his attention on worldwide wireless communication. He designed and manufactured a high-frequency generator of mechanical design that could produce large amounts of power at frequencies as high as 50 kilohertz. This design was later adopted by Ernst Alexanderson in producing radio transmission equipment for Marconi.

Moving his experimentation to Colorado Springs, Colorado, in 1899, Tesla constructed the world's largest radio frequency step-up transformer, which was self-resonant at about 50 kilohertz. When about 50 kilowatts of power was supplied to the primary winding of the transformer by the special high-frequency generator, a voltage of at least 12 million volts was induced in the secondary. The result was a tremendous electrical display—the first man-made lightning. At the same time an enormous electromagnetic field was generated that extended for miles and was strong enough at a distance of 26 miles from its source to light to full brilliance a 10 kilowatt bank of incandescent lamps connected to a resonant antenna. In the ensuing eighty years, no engineer has duplicated Tesla's feat.

Tesla later developed and patented the concepts of the tuned circuit, of resonance, and of use of the same resonant antenna for both transmission and reception. These ideas were to prove important in all forms of radio communication.

Ernst Alexanderson (1878–1970)

Educated in his native Sweden but fluent in many languages, Ernst Alexanderson had little difficulty in finding employment with General Electric when he came to the United States as a young man. Because of his background in electrical engineering, he was chosen to head a research and development group that would design and build a high-frequency generator for use in radio communication. This device, similar to the alternator in today's automobile, had been requested for a communications system that would use continuous radio waves (as modern broadcast stations do) rather than waves produced by an electric arc. The Alexanderson alternator (Figure 1.5), as the machine was called, developed a power of 2 kilowatts on a frequency of 100 kilohertz. Its first public demonstration came on Christmas Eve of 1906, when Reginald Fessenden, a professor of electrical

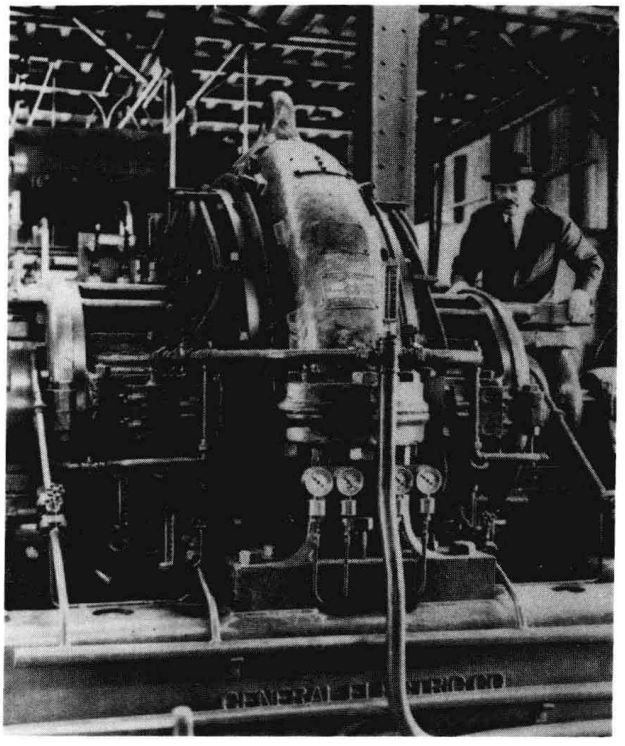


Figure 1.5 Alexanderson Alternator Dr. Ernst Alexanderson bridged the gap between Marconi's arc transmitter and De Forest's vacuum tube model. The Alexanderson alternator made possible the first broadcast of voice and music in 1906. (Photo courtesy of General Electric)

engineering at the University of Pittsburgh, broadcast both voice and music from an installation at his communications center at Brant Rock, Massachusetts. Well publicized in advance and heard over a wide area of land and sea, these first scheduled radio broadcasts of voice and music in the United States were made possible by the radio waves generated by the Alexanderson alternator.

Marconi soon heard of the new way of generating radio waves and ordered first a 50 kilowatt unit, then a 200 kilowatt unit, for his communication centers. RCA was quick to add several similar alternators ranging in power up to 200 kilowatts at their Radio Central, located at Rocky Point, Long Island. In the early 1900s, Lee De Forest invented the three-element (triode) vacuum tube (see Chapter 7), which was capable of generating continuous radio waves without high-speed rotating parts. The vacuum tube made the Alexanderson alternator technically obsolete because all future transmitters were designed to generate radio frequency waves by use of active devices like vacuum tubes and transistors.

Vladimir Zworykin (1889–1982)

Born in Russia, Vladimir Zworykin moved to the United States and found employment at Westinghouse Electric in the design of electronic equipment. He became particularly interested in the **cathode-ray tube,** in which electrons liberated from a heated cathode are formed into moving beams by electromagnetic or electrostatic fields and are then used to trace a pattern or picture on a phosphor-coated screen. Zworykin was not only an imaginative engineer but also an excellent artisan who constructed nearly all the equipment used in his experimentation. By 1928, he had perfected the cathode-ray tube to the point where he could take out a patent on it. In television the cathode-ray tube was important as both the "camera tube" that converts light patterns into a pulsating electrical current and the "picture tube" that converts the pulsating electrical current back into light patterns.

In 1929, Zworykin was appointed director of research at RCA and assigned to develop a completely electronic system of scanning for the television industry. The system had to be perfected and standards established and adopted by the Federal Communications Commission (FCC) before mass production of TV receivers and transmitting equipment could get underway (see Chapter 10). World War II delayed full-scale manufacture and distribution of black-and-white TV equipment until after 1945, although RCA had already invested many millions in TV research and development.

By 1948, Zworykin faced a new problem. The FCC had refused to accept RCA-developed television standards because they failed to take into account the growing interest in color television. It was not clear whether color telecasts could be satisfactorily received by owners of black-and-white TV sets. Therefore, the FCC froze the granting of licenses to new TV stations until the problem was resolved. By 1952, RCA had developed a new system of color television that could be satisfactorily received on the older monochrome receivers. The FCC freeze was lifted, and TV stations again received licenses without restriction.

John Bardeen, Walter Brattain, and William Shockley

Unlike many important breakthroughs in electronics, the transistor was not the result of a chance observation by a single struggling inventor. Its development was the result of a team effort begun in the mid-1940s by three highly skilled scientists working in the laboratories of the Bell Telephone Company: John Bardeen, Walter Brattain, and William Shockley. Each man had a slightly different background that allowed him special insights.

The Bell Laboratories research team (Figure 1.6) had as its goal the design and development of a new active device to take the place of the vacuum tube in performing both amplification and oscillation. An *active device*—for example, a transistor—is an electronic component that exhibits a large change in its output current when a small, changing voltage known as the *signal* is applied to its control element. The vacuum tube was a *thermionic device*—that is, it

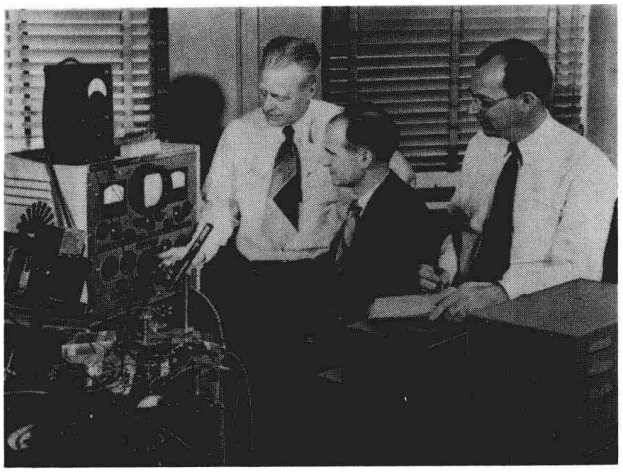


Figure 1.6 Bell Laboratories Research Team That Invented the Transistor Walter Brattain, William Shockley, and John Bardeen were awarded the Nobel Prize for physics in 1956. (Photo courtesy of the American Institute of Physics, Niels Bohr Collection)

depended on electrons liberated from a heated cathode for its operation. An inordinate amount of electrical energy was required to produce this heating, which not only set limits on how much the physical size of the tube could be reduced, but also eventually destroyed the tube. If in the new active device free electrons could be generated and liberated without the use of heat, the device's power consumption and physical size could be greatly reduced.

The family of substances known as **semiconductors** appeared to the Bell researchers as a promising area of study since these materials, normally nonconductors, could be made to conduct heavily when conditions were slightly changed.

Bardeen, Brattain, and Shockley initially used a thin sheet of semiconductor material (germanium) contacted by two springy pieces of phosphor-bronze wire. Some amplification (gain in signal strength) took place at close spacings, but it proved inadequate for practical application. The researchers then tried a different technique. A thin sheet of gold attached to an insulating base was cut in two with a razor blade, and a sheet of germanium was pressed firmly down upon it. This time somewhat more amplification was obtained, but still not enough to be practical. Announced publicly in 1948 as the point-contact transistor (Figure 1.7), the new device remained a laboratory curiosity.

Shockley then approached the problem from a different direction. Starting with germanium of the highest possible chemical purity, he formed it into thin sheets. Part of the sheets were treated with the vapor of an element such as arsenic, which resulted in germanium with more free electrons than usual, or N-type (negative) germanium. Other sheets of pure germanium were treated with the vapor of an element such as aluminum, which resulted in germanium deficient in free electrons, or P-type (positive) germanium. The elements used to treat the pure

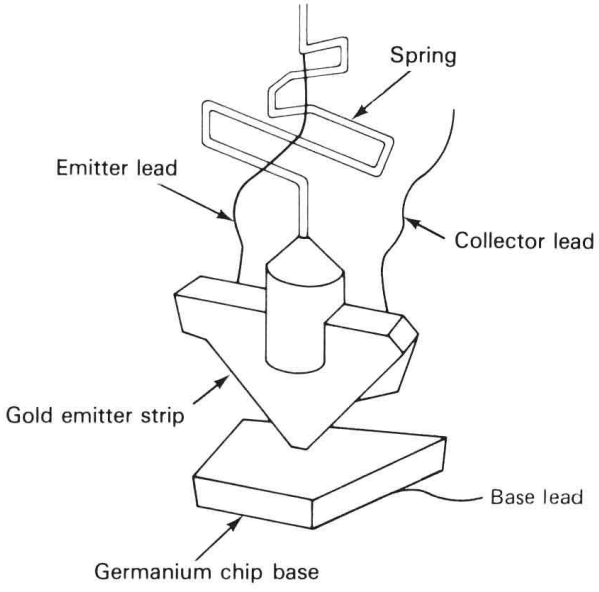


Figure 1.7 Point-Contact Transistor A spring forces a wedge of plastic, covered with gold leaf, against a slab of specially treated crystalline germanium. Produced in 1948, it was the first of a series of transistors created to overcome amplification problems.

germanium were called *dopants,* and it was only necessary to use about 1 atom of dopant to every 100 million atoms of the semiconductor.

When members of the Bell research team began assembling sheets of these materials into electrical "sandwiches," they realized they were on the right track at last. The sandwiches could be assembled in either of two patterns: PNP or NPN. Each configuration worked equally well; only the battery voltages needed to be reversed. After more research, development of the **bipolar transistor** (Figure 1.8) was announced. A fellow researcher, John R. Pierce, suggested the term *transistor* for the new device since it transferred an electrical signal through the germanium that would normally be a resistor. (Transistors are discussed in Chapter 9.)

The bipolar transistor developed rapidly into an active device capable of equaling or surpassing the performance of the previously unchallenged vacuum tube. The solid-state age had arrived. The transistor's extra-small size, coolness of operation, great life expectancy, and consistent reliability were to change completely the course of the Electronics Industry. For their contribution to science, Bardeen, Brattain, and Shockley were awarded the Nobel Prize in 1956.

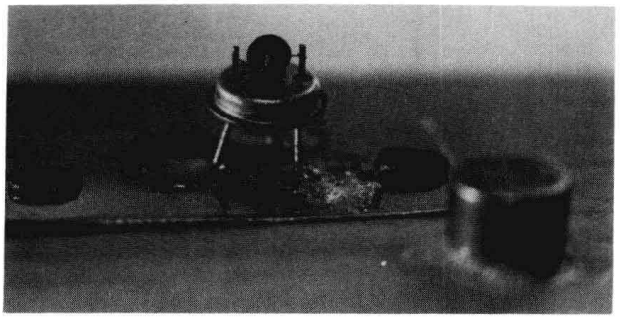


Figure 1.8 Bipolar Transistor This transistor, of the NPN type, is shown with its protective metal cap removed. The circular disc with two wires attached is the "electrical sandwich" of treated silicon.

Robert Noyce and Jack Kilby

Today's **integrated circuit** (IC) has been under development since the late 1960s, when researcher Robert Noyce of Fairchild Semiconductor and his counterpart at Texas Instruments, Jack Kilby, both working independently, started their experiments.

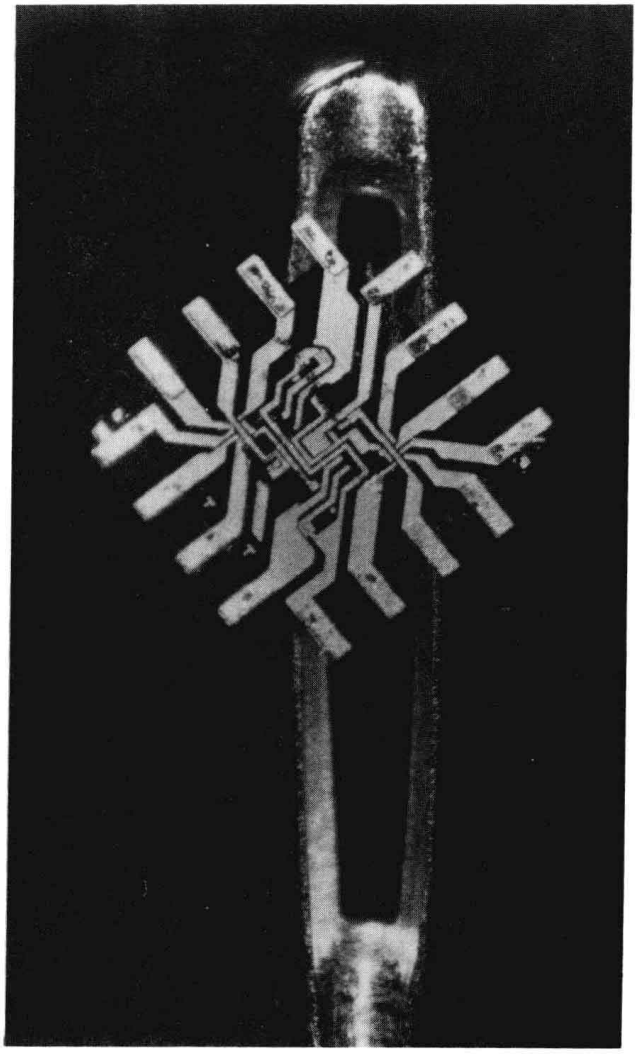


Figure 1.9 Early Version of an Integrated Circuit The IC has been placed over the eye of a needle to show its size. It probably contains five or six transistors. (Photo courtesy of Western Electric)

Their goal was to form a number of transistors, together with their associated components and wiring, directly on a tiny chip of silicon that served as a substrate, or base (Figure 1.9). They achieved their goal by treating one surface of the silicon with the high-temperature vapors of elements such as arsenic or boron. Their initial achievement—production of 10 to 12 separate transistors on a silicon chip—was considered remarkable. Today ICs with at least 100,000 transistors on their surface are widely used in mass-produced electronic equipment. (ICs are fully discussed in Chapter 7.)

SUMMARY

The Electronics Industry has moved from offering a few products that gross only several million dollars a year to offering a wide spectrum of products and services that produce at least $42 billion a year for its major divisions alone. The event that triggered this growth was the public's interest in radio broadcasting starting around 1922. Growth was so swift that radio receiver manufacturers were unable to keep abreast of the demand for new sets.

As new and better components and circuits were developed for radio receivers, experimenters were quick to adapt them to both related and new applications of electronics. Each new application seemed to open the way for further uses of electronics, and development has continued at an accelerated rate for six decades.

The term *Electronics Industry,* as defined by the Department of Commerce, covers a number of specialized divisions that provide unique products and services. The first major division, Electronic Equipment, is subdivided into Consumer Electronics, which provides products primarily intended for use in the home, and Electronic Systems, which turns out sophisticated electronic units for sale largely to government agencies and big business firms. The second major division, Electronic Components, represents production lines that deliver parts to the other divisions of the industry for assembly into salable products. Communications Electronics, one of the smaller divisions, comprises telephone, broadcasting, and television and employs far fewer people than the other divisions. Electronics Repair and Maintenance (appliance and TV repair) is so loosely organized that it defies close analysis.

A review of early developments in electricity and electronics shows that certain researchers made key contributions that influenced the thinking of their contemporaries and had a lasting impact on the course of today's electronics. Benjamin Franklin's great contribution was his view that electricity is a liquid that moves progressively from the positive to the negative pole of the circuit. Hans Christian Oersted established the relationship between electric current flow and electromagnetism. George Ohm, working with the most elementary equipment, established the definite relationship between voltage, resistance, and current flow that we know today as Ohm's Law. Joseph Henry, primarily interested in electromagnetism and its applications, served the science community well as secretary of the Smithsonian Institution and expedited development of both the telegraph and the telephone. Samuel Morse, more an artist than a technician, devised the telegraphic code that bears his name and conducted the first public demonstration of telegraphic communication.

Alexander Graham Bell, talented in many fields, was the first person to transmit the human voice over the telephone. Guglielmo Marconi, a shrewd businessman as well as a gifted experimenter, was the first to see commercial advantages in the assembly of electronic equipment for public sale and in sending and receiving wireless messages. Ernst Alexanderson broke with Hertz's and Marconi's tradition of transmitters powered with a crashing electric arc and turned to the generation of continuous radio waves—a feat that made the broadcasting of music and voice possible. Vladimir Zworykin, director of research at RCA during television's early development, resolved many complex problems to bring us color television.

By the 1950s electronics had become so complex and involved in the physical sciences that a breakthrough by an independent researcher had be-

come unlikely. The research team approach became a necessity, and the Bell Laboratories invention of the transistor was the result of a team effort. This approach changed the course of the industry. A decade later the combined skills of a score of researchers in a dozen electronics plants resulted in the breakthrough represented by the integrated circuit.

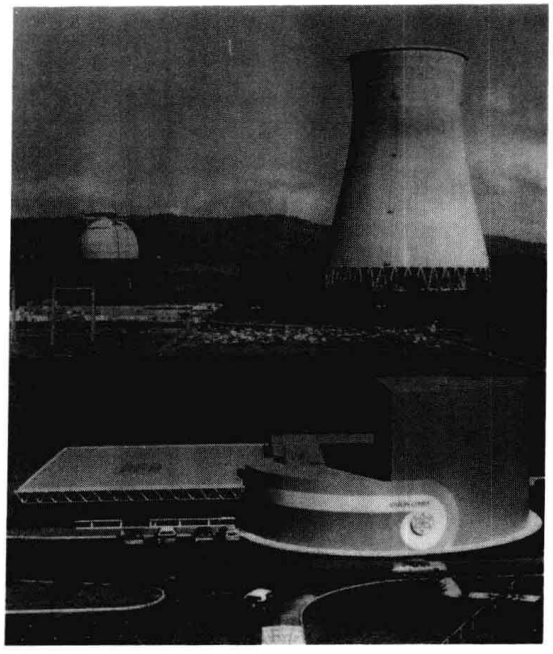
KEY WORDS AND CONCEPTS

alternator
bipolar transistor
cathode-ray tube
electricity
electronics
generator
integrated circuit
laser beam
ohm
resistance
semiconductor

SELF-TEST QUESTIONS

1. What key development in the early 1920s led to expanded use of electronics?
2. What are the two major divisions of the Electronics Industry?
3. Why is the manufacturing level for electronic parts and components quickly affected by cancellation of large government contracts for major electronic equipment items?
4. What are some of the employment benefits that electronics firms commonly offer new workers?
5. What are some of the measures used by West Coast electronics firms to increase the number of part-time employees?
6. How did Benjamin Franklin's theory of electrical current flow differ from the theory used in modern electronics?
7. What were Michael Faraday's three major breakthroughs in the field of electricity?
8. Although Joseph Henry added much to our knowledge of magnetism and electromagnetism, his greatest contribution was operating a clearing house of information for other inventors. What appointment did he hold that made this service possible?
9. What important invention of Edison and Berliner became a vital component of the modern telephone?
10. In what two areas of electronics did Marconi see commercial possibilities?
11. What well-known American firm was originally organized to take over the patents and other assets of the American Marconi Company?
12. How did George Westinghouse and Nikola Tesla help each other?
13. The Alexanderson alternator was quickly replaced as a generator of radio frequency waves by a purely electronic device. What was this device called?
14. What do electronics technicians call a specialized component that shows a large change in output current in response to a small, changing voltage applied to its control element?
15. What is an integrated circuit, or IC?

2 Basic Concepts and Definitions



The Trojan Nuclear Plant at St. Helens, Oregon, is the largest in the United States. The cooling tower for a heat exchanger is on the right. The domed building on the left contains the nuclear fission reactor and control equipment. (Photo courtesy of Portland General Electric)

Atomic Theory of Matter
- Elements
- Mass of Atoms
- Free Electrons
- Ions and the Generation of Electricity
- Fission
- Fusion
- Compounds

The Concept of an Electrical Circuit
- Current Flow in the Basic Circuit
- Electrical Power
- Ways of Generating Electrical Pressure

Summary

Key Words and Concepts

Self-Test Questions

Everything in the world about us is made up of **atoms,** which in turn are made up of wildly gyrating particles of negative and positive electricity. It's a little difficult at first to accept the idea that an alarm clock, the water in the shower, the clothes we wear, the food we eat, the newspaper we read, and the car we drive—even our own bodies—are all made up of uncountable numbers of atoms that are merely collections of negative and positive electrical charges. Yet the evidence is overwhelming that all matter consists of atoms.

ATOMIC THEORY OF MATTER

Atoms are arranged in many patterns, yet they all have a central nucleus of positive electricity made up of one or more **protons** (positive electrical charges) around which one or more **electrons** (negative particles) revolve at incredible speed and in well-defined orbits. In all normal atoms the number of positive charges in the nucleus is equal to the number of negative charges.

In many ways the structure of the atom is very similar to that of our solar system, whose planets revolve around a central sun. In the solar system the *centrifugal force* that tends to send the planets further into outer space is balanced by the *gravitational force* that exists between the planets and the sun and maintains the planets in relatively constant orbits. Centrifugal force is the same force you experience when you whirl a weight attached to a string. Centrifugal force also acts on the electrons of an atom but is balanced by the *attractive force* of the positive nucleus for the negative electrons (unlike charges attract). The attractive force holds the electrons firmly in their orbits.

Elements

All naturally occurring atoms must fall into one of 92 structural patterns based on the number of positive charges (protons) in the nucleus and the number and arrangement of the orbiting negative charges (electrons). A material made up of atoms whose internal structure is identical is known as an **element.** For example, if all of a substance's atoms have 29 electrons, we can be sure that they are atoms of the element copper; if the atoms all have 82 electrons, the element is lead.

Simplest of the atom communities is the element hydrogen. Hydrogen is a colorless, odorless gas whose atom has 1 proton (positive charge) in its nucleus, around which spins a single electron with a negative charge. The most complex structure is the radioactive metal uranium, with 92 protons in its nucleus, around which spin 92 electrons arranged in 7 orbital paths, or shells.

Spectroscopic analysis of distant stars and information gained from National Aeronautics and Space Administration (NASA) probes indicate that the 92 naturally occurring elements are distributed throughout our solar system and for hundreds of light years into outer space.

In the nuclei of the atoms of all elements except hydrogen are additional particles known as **neutrons,** which consist of a proton and an electron firmly bonded together. Since the positive charge of the proton is exactly canceled by the negative charge of the electron, a neutron has a net charge of zero. Therefore, neutrons are known as *neutral particles* of electricity. They have little effect other than to increase the mass of the nucleus.

Physical Characteristics of Elements

The number of electrons in the atoms of a certain element determines that element's physical characteristics. Hydrogen, with 1 electron and 1 proton, is a colorless, buoyant gas ideal for filling weather balloons. Helium, with 2 electrons and 2 protons, is also a colorless gas, but balloons filled with helium are less buoyant than those filled with hydrogen. Lithium, with 3 electrons and 3 protons, is a silvery, lustrous metal. Carbon, with 6 electrons and 6 protons, exists in its purest form as the diamond, and also as graphite, coke, and one of the principal ingredients of coal. At the other end of the spectrum is the silvery, radioactive metal uranium, which is appreciably heavier than lead.

Artificial Elements

By bombarding uranium with electric charges of the right type and velocity, it is possible to synthesize elements with as many as 103 electrons and protons. Perhaps the most important synthetic elements are the metals neptunium, with 93 electrons, and plutonium, with 94 electrons. Plutonium readily splits, or fissions, into less complex atoms and can be used to manufacture atomic bombs or to produce heat in nuclear power plants.

Mass of Atoms

It is convenient to think of an electron as the smallest possible charge of negative electricity. The proton, the positive charge, has about 1800 times the mass of the electron. Since neutrons are paired protons and electrons, they are about the same in mass as a proton. Surprisingly, atoms are mostly space. For example, if a uranium atom could be expanded in size until it was exactly one mile in diameter, its nucleus would only be about the size of a baseball.

Why do atoms give the impression of such great solidity when they are almost entirely space enclosing a few miniscule positive and negative electrical charges? The Danish scientist Niels Bohr, the man who first proposed the planetary theory of the atom's structure (Figure 2.1), came up with the answer. The "hardness" of atoms is due to the almost unbelievable speed of their orbiting electrons, which Bohr estimated at 7 million billion times per second. No wonder a piece of copper, nickel, or iron feels so dense and hard.

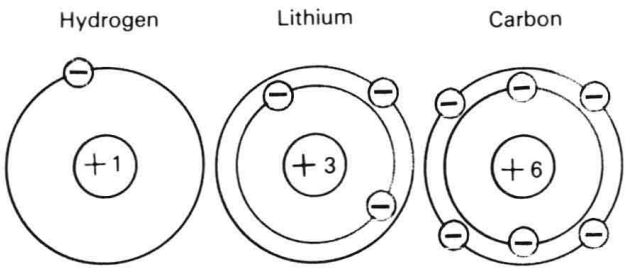


Figure 2.1 Planetary Theory of Atomic Structure Niels Bohr theorized that the various elements had atoms arranged like planetary systems.

Free Electrons

As previously noted, the electrons of atom communities rotate around the positive nucleus at very high speeds in well-defined orbits, or shells. The shell next to the nucleus can hold only 2 electrons; the second shell, 8; the third and all subsequent shells, 18 or more. The outermost shell, however, can have no more than 8 electrons but may have as few as 1. Metals that have from 1 to 3 electrons in the outermost shell are the best electrical conductors, or carriers of an electrical current.

Electrical conductor: A material, usually metallic wires, that readily permits the flow of an electrical current (electron flow). A conductor must provide an abundance of free electrons since an interchange of free electrons between atoms is essential for current flow to take place.

Apparently, when only a few electrons are in the outermost orbit, the strong attractive force that normally exists between the positive nucleus and the electrons with their negative charges is considerably diminished. Thus, an electron in the outermost shell can break away from the atom and move about as a *free electron*. Should enough free electrons break away from enough atoms and move progressively from the negative to the positive pole of the circuit, we can say an *electrical current* is flowing.

Silver, with one electron in its outer shell, is the best electrical conductor but is too expensive for most commercial applications. In descending order of conductivity (the ability to carry an electrical current with ease) are the elements silver, copper, gold,

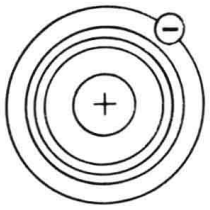


Figure 2.2 A Copper Atom Copper atoms contain 29 electrons arranged in 4 shells, with only 1 loosely held electron in the outermost shell. Therefore, copper is a good electrical conductor.

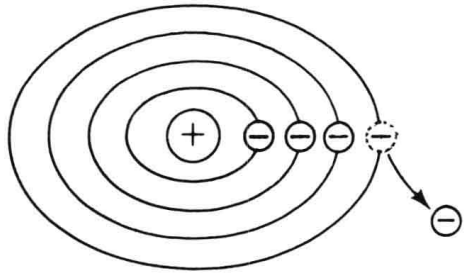


Figure 2.3 An Ion If an atom's outermost electron is freed, the atom becomes an ion and is no longer electrically neutral.

and aluminum. Iron is far down the list. Copper (Figure 2.2), because of its high conductivity (or low resistance) and reasonable cost, is the most widely used conductor, although aluminum has recently been used for applications such as house wiring. However, the difficulty in making low-resistance connections when aluminum is used is a continuing one since the use of solder with aluminum is impractical.

Resistance: The opposition a substance offers to the flow of an electrical current.

Ions and the Generation of Electricity

Ion: An atom that is no longer electrically neutral.

As previously noted, in all normal atoms the number of positive charges in the nucleus is exactly equal to the number of negative charges represented by the orbital electrons. Consequently, the atom's normal charge is zero—in other words, it is electrically neutral. But should an atom of copper, for example, lose an electron from its outermost orbit as a free electron, the atom would no longer be electrically neutral. It would have a net charge of +1 (Figure 2.3) and would become an **ion.** In other words, the difference between a copper atom and a copper ion is that the first is neutral while the second carries a positive charge. Because all electrons of all elements are exactly alike, there is always a tendency for a copper ion to capture a passing electron and once more become a neutral atom.

When all the atoms in a particular circuit are electrically neutral (the normal state of affairs) and no free electrons are present, there is no electrical activity. To produce electrical activity, we must deliberately tear electrons away from the outermost shell of each atom to create both positive ions and negative free electrons. Electrical activity will continue until all the ions have captured all the free electrons and have once more become neutral atoms. Electrical activity then ceases.

The process of generating electricity, then, is basically that of tearing electrons away from atoms and storing the free electrons at the negative pole of the circuit and the ions at the positive pole. Methods of generating electrical activity are described later in the chapter.

Fission

Fission: The process by which very large atoms (such as uranium or plutonium atoms) are split into smaller ones.

All nuclear reactors currently in operation are of the **fission** type. During fission, enormous amounts of heat are liberated. This heat generates high-pressure steam, which is used to spin the shafts of very large AC generators. A single nuclear plant can produce as much electricity as a major hydroelectric installation. The fission process itself does not produce useful electricity. The tremendous amount of heat that fission generates does the useful work. The AC generating equipment used is exactly the same type used at any oil- or coal-fired power generating plant. The principal disadvantages of the fission method of power generation are the comparatively large amount of radioactive waste that must be carefully stored, the large amount of relatively scarce uranium required, and the comparatively low efficiency of the fission process. Danger is also posed by the possibility of an uncontrolled nuclear reaction.

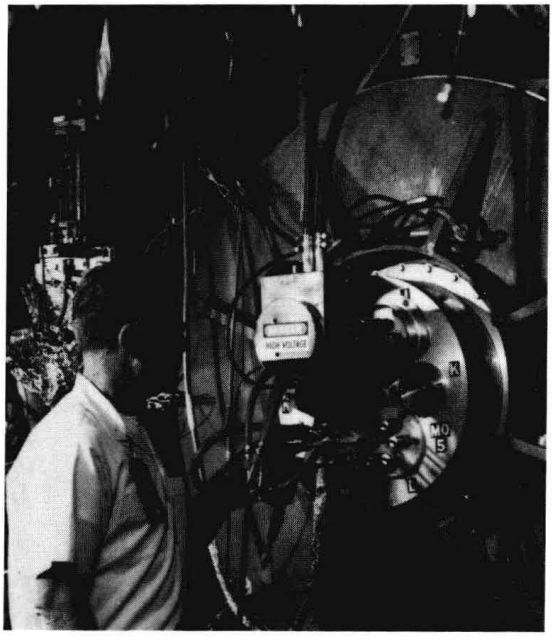
Fusion

Fusion: The process by which very simple elements combine to form more complex ones.

Large amounts of energy can also be derived by the process of **fusion.** The sun's heat is generated by the fusion, or combination, of hydrogen atoms to form helium atoms. Extremely high temperatures and pressures are required; such conditions are readily found on the sun's surface. The fusion process on the sun provides heat and light for our entire solar system, and astronomers believe fusion is responsible for the light given off by all the stars in the reaches of outer space.

We have been able to bring about atomic fusion for only the briefest time on the surface of the earth—the time necessary to explode a hydrogen bomb. To provide the heat and pressure needed to initiate the fusion process, it is necessary to use a conventional fission-type atomic bomb as a "trigger." Research on a way to accomplish controlled fusion is currently under way, focusing at present on the use of immensely strong magnetic fields.

Were the atomic fusion process to be developed to the point where it is both safe and practical, it would offer some advantages over nuclear fission. Hydrogen is both abundant and cheap, the energy yield of fusion is very high, and few radioactive



The nuclear fusion device "Alice," installed at the Lawrence Radiation Laboratory, is one of four experimental projects in the field of fusion being conducted by the USDOE. Here, a technician secures the fusion chamber port. (Photo courtesy of USDOE)

byproducts are formed. On the other hand, the number of conventional nuclear plants of the fission type is slowly growing, and they do offer a reasonable alternative to oil- or coal-fired generating plants. A golf-ball-sized lump of uranium U-235 (containing more neutrons than ordinary uranium) will produce more heat than will 2 million pounds of coal.

Compounds

It is technically correct, but not quite inclusive enough, to say that everything is made up of elements, which are in turn made of electrons and protons. We should supplement this statement with the observation that elements frequently combine to form **compounds.** For example, an atom of sodium is combined with an atom of chlorine to form sodium chloride, or ordinary table salt. Carbon, hydrogen, and oxygen are combined to form complex compounds like sugar or carbohydrates. Compounds, however, can always be broken down again, and we can recover the original elements and their atoms.

THE CONCEPT OF AN ELECTRICAL CIRCUIT

A **circuit** is a continuous path of wires or circuit elements over which current can flow. All electrical circuits, no matter how complex, can be broken down into a single basic circuit with three fundamental components (Figure 2.4). The *source* is the origin of the circuit's energy. Normally its polarity is shown, being marked with conventional + and − signs. The positive pole of the source represents atoms that have lost electrons and become ions. The negative pole represents a concentration of free electrons that have been torn away from atoms. While the symbol for a generator is shown in Figure 2.4, the source could just as well be a battery, a photoelectric cell, a vacuum tube, or a transistor.

The second basic component of the circuit is the *load,* which converts the energy supplied by the source into other forms of energy such as heat, light, and motion. Figure 2.4 shows a lamp filament as the load, but the load could just as well be a resistor, a motor, a loudspeaker, or even a radio or a television set.

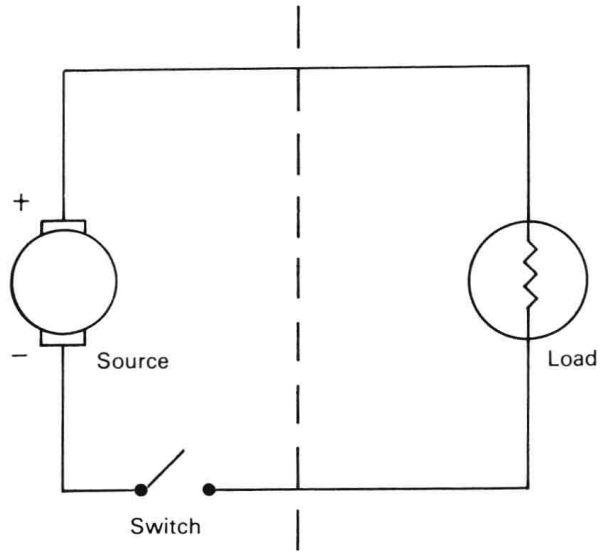


Figure 2.4 A Basic Electrical Circuit The circuit consists of a source, a load, and a switch.

The third component, the switch, is made up of metallic wires that complete the circuit between the terminals on the source and the load. They are called *conductors* since they readily permit the flow of electrons from the negative pole of the source to the load, then from the opposite end of the load back to the source. Without conductors the circuit would be open, or incomplete, and no current could flow.

Good conductors invariably offer low resistance, or opposition, to current flow. Metals such as silver, copper, and aluminum are widely used as electrical conductors because of their very low resistance to current flow. Electrons in the outermost shells of these elements' atoms interchange freely with similarly placed electrons of adjacent atoms, even at room temperatures. All electrons are identical and can interchange freely with any other electrons in the outermost orbit of any atom. Under

proper circuit conditions, this random interchange of free electrons can become an **electrical current**—that is, a continuing drift of electrons along the conductor from the negative to the positive pole of the circuit.

Current Flow in the Basic Circuit

For current to flow in an electrical circuit, certain requirements must be met. The circuit must be complete, or continuous, with both terminals of the source connected to both terminals of the load through suitable low-resistance conductors. There must be enough electrical pressure to drive sufficient numbers of electrons through the circuit and to overcome its resistance to their flow. Formerly called *electromotive force* (EMF), the electrical pressure required is now referred to as *voltage*.

For voltage to exist, electrons must be separated from the atoms of which they are a part and deposited at the negative pole of the circuit. The atoms from which the electrons have been torn away are now ions, or atoms with a positive charge, and they are deposited at the positive pole of the circuit. In other words, the negative pole of the circuit represents an excess of free electrons. Conversely, the positive pole represents a deficiency of free electrons. The relative degree of electron excess and deficiency is expressed in terms of voltage. The strength of the current that flows is expressed in **amperes.**

In the circuit shown in Figure 2.4, when the switch is closed, a progressive movement of electrons immediately takes place from the negative pole of the source through the load and back to the positive pole of the source, where the electrons start combining with positively charged ions to form neutral atoms once more. The driving force that moves the electrons through the circuit is the opposition of like electrical charges to each other. The electrons, all carrying a negative charge, are forced away from the negative pole (where they are in excess) toward the point of maximum electron deficiency at the positive pole. Students often ask why the positive ions don't move too. The answer is that the electrons are just a fraction of the size of the ions and are extremely mobile, so the electrons move more readily when the current flows.

Battery as a Source

If the source in a circuit is a battery, eventually all electrons in excess at the negative pole will move through the load to the positive pole of the battery, recombining with the positive ions there to form neutral atoms. When there is no longer any excess of electrons at the negative pole or any deficiency of electrons at the positive pole, we say that the battery is run down, or discharged.

With a conventional automotive battery, we can apply a DC voltage across the battery terminals from the alternator's output. The voltage once more tears electrons away from atoms of the material making up the battery plates. The electrons are deposited at the negative terminal of the circuit, and the positive ions are deposited at the positive terminal. Current can then be supplied once more by the battery until the relative excess and deficiency of electrons again disappears. In the case of a flashlight cell, recharging is impractical; we merely discard the run-down unit and replace it.

Electron Motion

Free electrons, move through the circuit in a somewhat random manner, moving from one atom of the conductor to the next at a comparatively slow rate—perhaps averaging several inches per minute. But the wave motion, or impulse, that their atom-to-atom movement initiates takes place at the speed of light—that is, at about 186,300 miles per second.

The following analogy helps to explain electron flow. A 1 inch pipe extends from New York to San Francisco. It is completely filled with 1 inch ball bearings. When a technician in New York forces one more ball bearing into the pipe, another ball bearing is immediately forced out in San Francisco, 3000 miles away (Figure 2.5). Yet no individual ball bearing moved more than an inch. Similarly, while elec-

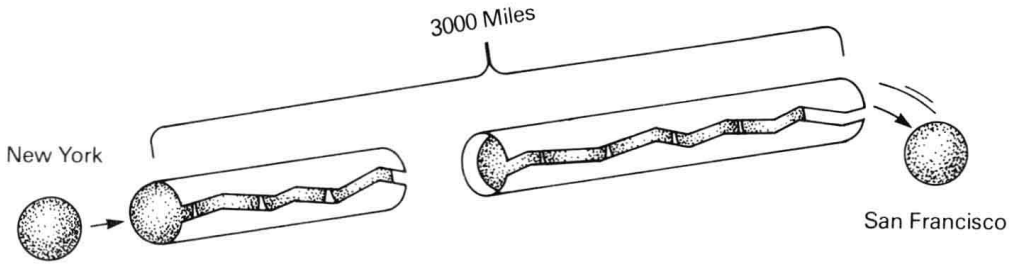


Figure 2.5 Electron Flow Analogy A 3000 mile chain of movement is initiated by an extra ball bearing added to a pipe filled with ball bearings. In the same way, movement of electrons along a conductor sets up a wave motion that travels at the speed of light.

trons move only a very small distance in a conductor, their movement triggers a wave motion that travels at the speed of light.

Conventional Theory of Current Flow

Benjamin Franklin's theory that electricity is a fluid that always moves from the positive to the negative pole of the circuit is the basis for the conventional, or plus-to-minus concept, of electrical current flow. Electrical engineers and journeymen electricians found the positive-to-negative theory ideal for circuit analysis. They continue to use the concept today. The schematic diagrams of both electrical and electronic equipment are also based on the same premise: that current moves from the point of highest positive potential through the circuit and eventually to ground (usually the negative point of the circuit). The conventional theory should be used whenever appropriate.

Electron Theory of Current Flow

When engineers began to design radio receiving and transmitting equipment on a large scale in the 1920s, they found they could not explain the operation of the vacuum tube in terms of conventional current theory. They therefore adopted the *electron theory*—the theory that the flow of electrical current is caused by the progressive movement of electrons along a conductor from the negative to the positive pole of the circuit.

Six decades later it is still easier to explain the operation of active devices, such as field effect transistors, in terms of the electron theory, rather than the conventional theory. The electron theory is generally accepted by the Electronics Industry, and FCC operator examinations as well as job placement exams for government agencies are based upon it. This book uses the electron theory throughout.

Electrical Power

The electrical power in a circuit is computed by multiplying the electrical pressure applied across the circuit (measured in volts) by the current intensity through the circuit (measured in amperes). The power in the answer is measured in watts.

For a large amount of power to be generated or controlled in the equation

$$\text{watts} = \text{voltage} \times \text{amperage}$$

the amperage must be a substantial value so that the product of amperage and voltage will also be substantial. If the voltage is 10,000 volts and the amperage is zero, the power will also be zero, and no useful work can be performed. If a substantial amount of current is flowing, radio and TV transmitters can develop thousands of watts of power, the tubes or

Watt: The power expended when 1 volt drives a current of 1 ampere through a circuit having a resistance of 1 ohm.

Volt: The unit of electrical pressure; the amount of pressure required to drive a current having a strength of 1 ampere through a circuit having a resistance of 1 ohm. (The effective value of electrical pressure is the difference in potential voltage between the two points under study in the circuit.)

Ampere: The unit of current intensity; the strength of current flow that takes place when a pressure of 1 volt is applied across a circuit having a resistance of 1 ohm. (An ampere is actually a measure of how many electrons are in motion—that is, 6×10^{18} electrons per second.)

Ohm: The unit of electrical resistance; a circuit has a resistance of 1 ohm when an electrical pressure of 1 volt applied across it results in a current strength of 1 ampere through it.

transistors in the output stage of a stereo amplifier can deliver 50 to 100 watts of audio power to loudspeakers, and microwave ovens rated at 1500 watts can roast a turkey in record time.

Ways of Generating Electrical Pressure

Maintenance of electron flow in a circuit is dependent upon maintenance of an excess of free electrons at the negative pole of the circuit and a deficiency of free electrons at the positive pole. The following subsections discuss common methods of providing such a difference in potential, as well as the comparative economic importance and advantages and disadvantages of each method.

Historically, friction was the first technique used in developing a voltage. In the mid-18th century the chemical approach was used when Alessandro Volta developed the first cell and battery. Generation of pressure was revolutionized in the early 19th century by Michael Faraday's invention of the first electric generator. Other inventors later improved this basic device. Thomas Edison brought the storage battery and the DC generator to a high degree of efficiency; George Westinghouse and Nikola Tesla developed the AC generator and AC distribution systems in wide use today; and still other scientists made important contributions in such fields as thermal, photoelectric, and piezoelectric generation of electricity.

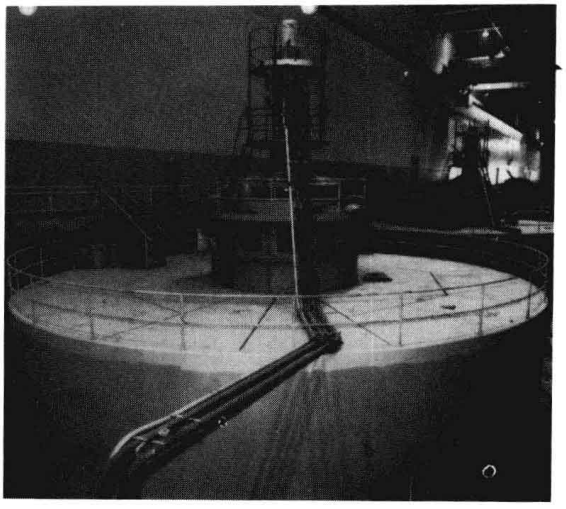
Electromagnetism

More than 90% of the world's electrical power is generated by rotating a coil of wire at high speed in a powerful electromagnetic field. In electromagnetism (see Chapter 4), the rotation of a coil of copper wire in a magnetic field pulls electrons away from the atoms of the copper and converts them to copper ions. This action provides the electrical potential that drives current through the circuit. The liberated electrons are called free electrons.

To produce large amounts of electrical power, very large amounts of mechanical energy are required. Such energy is normally provided by hydroelectric turbines or steam turbines. The steam turbines are powered by high-pressure steam produced either by the burning of fossil fuels (oil and coal) or by atomic fission. The machine used to generate alternating current is called an *alternator*, or less commonly, an *AC generator*. If atomic fusion ever proves to be a practical source of energy, it too will be used to provide high-pressure steam to spin the shaft of an alternator or AC generator.

Chemical Reactions

Next in importance in the production of the world's electricity are chemical reactions of a wide variety. Chemical reactions strip electrons away from



This BPA installation at The Dalles, Oregon, is typical of modern hydroelectric plants. Water-driven turbines spin powerful electromagnets past large fixed coils in the base to produce large amounts of electrical power. The stairs provide a scale for the size of the turbines. (Photo courtesy of Bonneville Power Administration, USDOE)

atoms and deposit the electrons at the negative pole of a battery. The positive ions formed collect at the positive pole.

Batteries are made up of two dissimilar metals (or dissimilar forms of the same metal) in contact with an **electrolyte**—a conductor that can be an acid, a base, or a salt dissolved in water. Primary cells are cells that cannot be recharged and must be discharged when worn out; secondary cells are cells that can be recharged through use of a DC source to reform the battery. In newer automotive batteries the electrolyte is sealed (but vented) in the cells, and the addition of distilled water is rarely required. Hydrogen gas, a product of the charging process, can escape, but most water vapor is condensed and returned to the cell.

Light

When small particles of light energy known as *photons* strike the atoms of elements such as cesium, silver, and silicon, the impact dislodges electrons from their outermost orbit and converts the atoms that have lost electrons into positive ions. Producing electrical energy from the sun's rays has recently aroused much interest, but the attempt can be very expensive. One mountaintop microwave transmitter installation required over $12,000 worth of silicon cells arranged in panels to provide 50 watts of power. In the case of satellites in outer space, high costs are easily justified, and many solar panels are used to charge batteries that, in turn, power a satellite's electronic circuits.

Heat

If two lengths of wire of dissimilar metals (e.g., iron and copper) are twisted together at one end and the twisted part heated, a voltage will appear across the open ends that is proportional to the difference in temperature between the heated end and the cool end. Such a device, called a *thermocouple,* can be a convenient source of electricity at remote locations. If one thermocouple does not supply enough voltage, two or more can be connected in series to provide the required voltage.

Friction

Friction was the first observed way of creating a difference in potential, or voltage. Friction may be produced in diverse ways—for example, the shuffling of shoes across a nylon carpet or the rubbing together of rising and falling air currents during a thunderstorm. Although a lightning flash may represent several million volts and millions of amperes, there is no practical way to harness this display of energy. In electronics we are often more concerned with minimizing friction-produced electricity than with producing it. For example, devices must be installed on an airplane's wing to dissipate the electrical charges produced by rapid movement of the wing through the atmosphere.

Piezoelectric Activity

A family of chemical compounds called *piezoelectric materials* have important electronic ap-

plications. (The Greek word *piezo* means "to squeeze.") When subjected to mechanical stress, electrical charges appear on the crystalline surfaces of these compounds. Conversely, when electrical charges are applied to their crystals, mechanical movement is produced. The compounds must be in crystalline form for this behavior to take place.

While piezoelectric activity has been recognized for at least a century, it was not applied commercially until the mid-1930s. At that time the first microphones and phonograph pickup cartridges were offered for sale. They employed thin slices of the crystalline form of potassium tartrate and were known as crystal microphones and crystal pickups. (Today's manufacturers prefer to use the double salt, barium titanate.)

Also in the 1930s, new standards for frequency stability of radio station signals forced manufacturers to use thin slices of quartz (silicon dioxide) in a special crystal-oscillator circuit to control the operating frequency of the transmitters they produced.

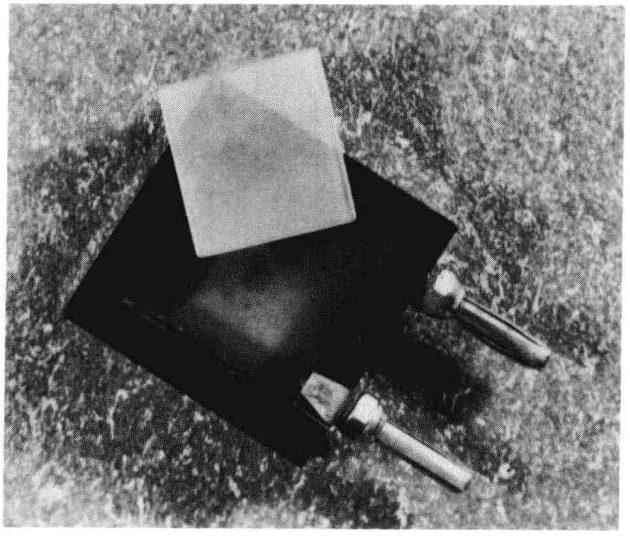


Figure 2.6 A Quartz Crystal and Its Holder from a Radio Transmitter
When stressed mechanically, electrical charges appear on the crystalline surface. Similar crystals are used in computers and digital clocks and watches.

All present-day AM, FM, and TV stations are kept on their assigned frequency by specially ground quartz crystals (Figure 2.6).

SUMMARY

The world about us is made up of countless atoms, the building blocks of all matter. Each atom is composed of a nucleus with a positive charge around which negative charges, called electrons, revolve at inconceivable speeds. In the normal atom, the number of positive charges (protons) in the nucleus is exactly equal to the number of orbiting electrons. Therefore, the atom has a net charge of zero and exhibits no electrical activity.

To generate electrical activity, we have to tear electrons away from atoms and store them at the negative pole of the circuit as free electrons. The atoms that have lost an electron are no longer neutral; they carry a positive charge and are known as ions. Ions are found at the positive pole of the circuit. Electrical activity continues until all free electrons are reunited with all the ions to form neutral atoms once more.

Throughout our solar system atoms are organized in 92 patterns. When a substance is made up entirely of atoms having an identical number of protons and electrons, it is said to be an element. Elements range from hydrogen, with 1 proton in its nucleus and 1 orbiting electron, to uranium, with 92 protons in its nucleus and 92 orbiting electrons.

Tearing electrons away from atom communities enables us to light our buildings; to power our industries; to light, heat, and cool our homes; and to provide communications for news and entertainment. Most electricity is generated by spinning coils of copper wire in intense magnetic fields. Chemically produced electricity makes up a comparatively small share of the total amount generated.

An electrical current is caused by a progressive movement of free electrons from the negative to the positive pole of the circuit. Although electrons move at a comparatively slow rate from atom to atom in the conducting material, the wave motion they gen-

erate travels at the speed of light—that is, at about 186,300 miles per second.

All electrons and protons of all atoms are exactly alike. Any atom that has lost an electron and temporarily becomes a positive ion can capture any electron from any source and immediately convert itself back to a neutral atom. This process is continuously under way.

When a pressure of 1 volt drives a current of 1 ampere through a circuit having a resistance of 1 ohm, 1 watt of power (a very small amount) is consumed. The power consumed by household appliances is much greater and is usually measured in kilowatts (1000 watts), while electrical power generating plants have their output expressed in megawatts (1,000,000 watts).

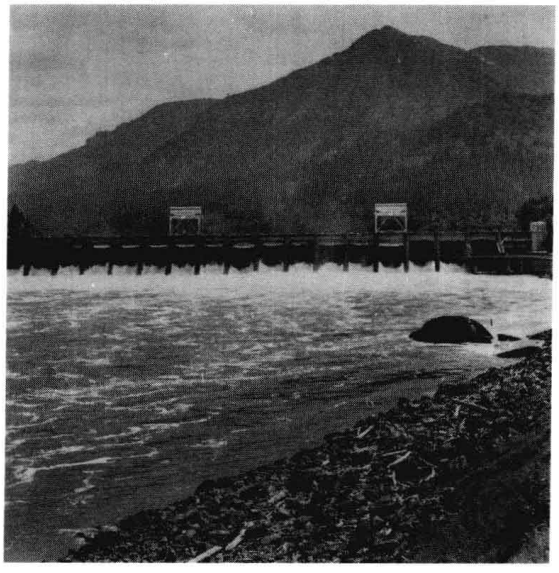
KEY WORDS AND CONCEPTS

ampere	electrolyte	neutron
atom	electron	ohm
circuit	element	proton
compound	fission	resistance
electrical conductor	fusion	volt
electrical current	ion	watt

SELF-TEST QUESTIONS

1. All atoms of the 92 naturally occurring elements are made up of three basic types of building blocks. What are they called?
2. Do elements such as silver or copper, which are good conductors of electrical current, have a large or a small number of electrons in the outermost orbit, or shell, of their atoms?
3. What is the difference between a copper atom and a copper ion?
4. Why is it necessary to tear electrons away from atoms to produce electrical activity?
5. Do good conductors offer low or high resistance to electron flow?
6. Does the negative pole of a battery represent a deficiency or an excess of free electrons?
7. How does the rate of movement of free electrons along a conductor compare with the speed of the wave motion that the electron movement creates?
8. In what fields of electricity and electronics might the conventional (plus-to-minus) theory of electrical current flow be used to advantage?
9. If you were studying for your amateur radio operator's license, why would you need to be familiar with the electron theory of current flow?
10. What are the two most important ways of producing large quantities of electrical power?
11. In a generator, or alternator, what is the source of the free electrons available at the terminals?
12. Why is electrical power produced by sunlight falling on photoelectric cells not in wide commercial use?

3 Basic Laws, Circuits, and Devices



The Bonneville Dam bridges the Columbia River just east of Portland, Oregon. The recent addition of powerful generators has increased capacity from 544 to 1062 megawatts. (Photo courtesy of Bonneville Power Administration, USDOE)

Coulomb's Law of Charges

Kirchhoff's Laws

Ohm's Law
- Three Basic Statements of Ohm's Law
- Problems and Solutions

Power Computations
- Power and Wattage
- The Time Factor
- Three Basic Formulas for Computing Power

Electrical Circuits
- Open and Short Circuits
- Series Circuits
- Parallel Circuits
- Series-Parallel Circuits

Resistance and Resistors
- Resistive Materials
- Carbon Resistors
- Wire-Wound Resistors

Variable Resistors
Color Coding Resistors

Summary

Key Words and Concepts

Self-Test Questions

Charles Coulomb, Gustav Kirchhoff, and George Ohm all formulated laws, or statements, relating to their electrical discoveries. Coulomb's Law of Charges is already familiar to us since we described its application in Chapter 2. Kirchhoff's Law of Current and Law of Voltage are still of use today in the evaluation of complex circuits, but Ohm's Law is perhaps the most important statement in the field of electronics and electricity.

COULOMB'S LAW OF CHARGES

Coulomb's Law of Charges: Like electrical charges repel each other, while unlike charges always attract.

Coulomb's Law was described in two instances in Chapter 2. The first was the powerful attractive force between the orbital electrons on an atom and the positive charge on its nucleus—a force that just balanced the centrifugal force caused by the electrons' extremely rapid rotation. The second was the force that moved electrons progressively along a conductor—a force we described as the "dislike" of free electrons accumulated at the negative pole of the circuit for other electrons at the same point. This opposition moved the electrons away from the negative pole, down the conductor, through the load, and to the positive pole of the circuit. Figure 3.1 illustrates Coulomb's Law of Charges.

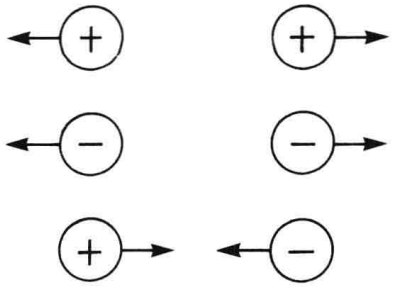


Figure 3.1 Coulomb's Law of Charges Like charges repel; unlike charges attract.

Later we will learn that Coulomb's Law is vital in the operation of electronic devices such as the precipitron, used to remove soot and ash from factory smokestacks, and the vacuum tube, where electrons liberated from the heated cathode are attracted to the tube's positive plate.

Coulomb later modified the original statement to read more precisely as follows: *The force of attraction or repulsion between two electrically charged bodies varies directly in proportion to the product of their charges and in inverse proportion to the square of the distance between the charges.*

For our purpose we can simplify this statement and make it easier to understand without affecting its validity. First, any time we have two *similar* electrical charges (either two positive or two negative charges) an *opposing* force exists between them that tends to move them farther apart. But if we have two *unlike* charges (one positive and one negative), an *attractive* force exists between them that tends to draw them together. The stronger the electrical charges and the closer their spacing, the greater the repulsion (opposition) or attraction between them.

Later Coulomb was to discover that magnetic poles follow the same basic rules of behavior as electrical charges: Like magnetic poles (two north poles or two south poles) repel each other, while a north and a south pole always attract each other (Figure 3.2). As with electrical charges, the degree of attraction or opposition between magnetic poles varies *in inverse proportion to the distance between them*. In other words, as you move a pair of poles

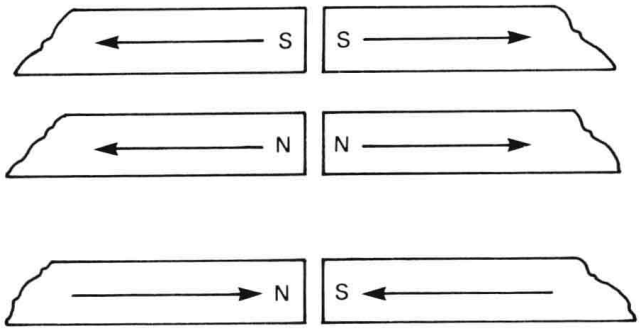


Figure 3.2 Coulomb's Law of Magnetic Poles Like poles repel; unlike poles attract.

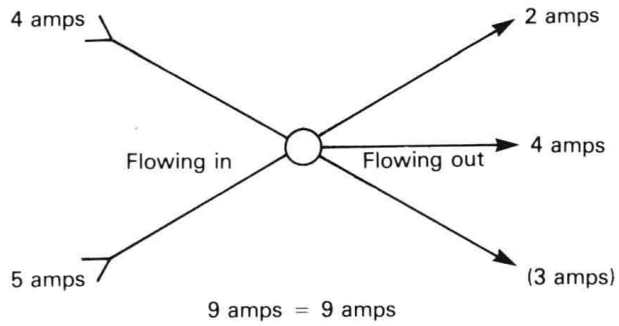


Figure 3.3 Kirchhoff's Law of Current The sum of the currents flowing into a junction must equal the sum of the outflowing currents.

away from each other, the attractive (or repulsive) force falls off very rapidly. As you move the same pair of poles closer to each other, the attractive (or repulsive) force increases very rapidly.

KIRCHHOFF'S LAWS

The German physicist Gustav Kirchhoff is principally known for his laboratory work in the spectroscopic analysis of materials, but he was a man of many interests. He was much impressed by Ohm's statement of the relationships between current, voltage, and resistance in a circuit (discussed in the next section). From Ohm's observation, Kirchhoff developed two laws of his own: one relating to current flow (known as the *point,* or *current,* law) and the other relating to what he called *voltage drop* around a circuit caused by the current flow.

> **Kirchhoff's Law of Current:** When a number of currents enter a junction (or a point in a circuit), their sum is equal to that of all the currents flowing away from that junction.

In Figure 3.3, currents of 4 amperes and 5 amperes are shown entering the junction, and currents of 2 amperes and 4 amperes are shown leaving the same junction. To determine the unknown remaining current (shown at the lower right) flowing away from the branch circuit, add the inflowing currents (4 amperes + 5 amperes, or 9 amperes), and from 9 subtract the sum of the known outflowing currents (2 amperes + 4 amperes, or 6 amperes). The unknown current through the branch circuit then must be 9 − 6, or 3 amperes.

Kirchhoff's Law is important because it establishes the principle that electricity, like water, is incompressible—that is, it is impossible for either more or less electrical current to flow into a junction than flows away from the same junction. To continue the water analogy: Assume that Figure 3.3 represents a hydraulic circuit with inflowing pipes bringing water to a junction such as a water meter. The circuit also contains outflowing pipes. The rate of flow in one incoming pipe is 4 gallons in a given length of time, while another pipe carries in 5 gallons. If, on the outgoing side, 2 gallons are flowing through one output pipe and 4 gallons through the second, the third unmetered pipe must be carrying 3 gallons of water. In other words, no more (or less) water can flow into a junction in a hydraulic circuit than flows away from the same junction. The amount of water flowing in must be exactly equal to the amount of water flowing out.

Electronics technicians make constant use of Kirchhoff's Law of Current. However complex the

circuit under study, technicians know that the amount of current flowing into the circuit has to be exactly equal to the amount of current flowing out of the same circuit.

> **Kirchhoff's Law of Voltage:** The sum of all the individual voltage drops around a circuit is equal to the voltage of the source.

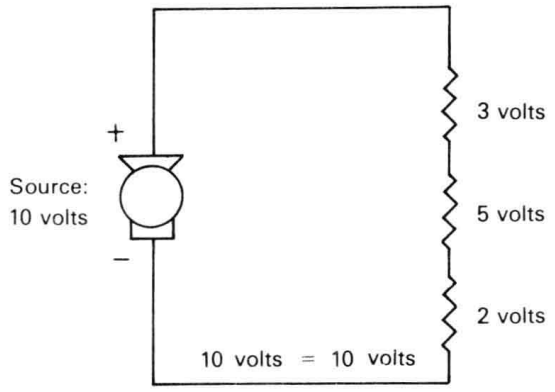


Figure 3.4 Kirchhoff's Law of Voltage The sum of voltage drops around a circuit equals the voltage at the source.

Kirchhoff's voltage law, also known as the *loop law,* is based on the following concept: As electrical pressure, or voltage, drives a progressive flow of electrons through the circuit (from the negative pole, through the load, and to the positive pole), all the voltage energy is eventually converted into energy forms such as heat and light so that finally no voltage energy is left. The voltage energy disappears because, as the electrons are driven through the load resistance by electrical pressure (voltage), a large number of collisions take place between the atoms of the resistive material and the moving electrons. Each collision changes some of the voltage energy into heat energy, which lessens the amount of voltage available. Eventually, as the electrons near the positive pole of the circuit, all the voltage has been used up by being converted to heat energy. But at that point there is no further resistance to overcome that would require the application of voltage, so electrons continue in their movement to the positive pole of the circuit.

Figure 3.4 illustrates how Kirchhoff's voltage law works. The load of the circuit has been divided into three resistors of unequal value that are connected in series. A **series circuit** is one that has a single path over which current flow can take place. Electrons start moving through the resistor nearest the negative pole of the circuit under the urging of 10 volts of pressure supplied by the generator. Voltage energy is converted into heat energy, and eventually 2 volts are used up. Therefore, 8 volts of energy are left at the junction of the lower and the middle resistor. As electrons move through the second resistor, additional voltage energy is converted to heat energy. At the junction of the second and the third resistor, only 3 volts of energy are left. As the electron flow continues through the third resistor, all the remaining 3 volts are used up. But there is no resistance to be overcome between that point and the positive pole of the generator; the electrons keep flowing over a very low-resistance conducting wire.

Kirchhoff referred to the change in voltage across a resistive element in a circuit as **voltage drop.** According to the voltage law, any time a source is connected to a load, the load's resistance converts all the electrical energy into heat energy. If the load is split into a number of parts (as illustrated in Figure 3.4), each part will produce a voltage drop in direct proportion to its resistance. If the voltage drops are added together, they will equal the generator's output voltage. (If the resistors are identical in value, the voltage drop across each resistor will be identical as well.)

Commercial applications of Kirchhoff's voltage law can probably be found in most homes. For many years millions of radio receivers, called AC/DC receivers (Figure 3.5), were manufactured with tube heaters connected in series. The voltage drop caused by each heater added up to the line voltage.

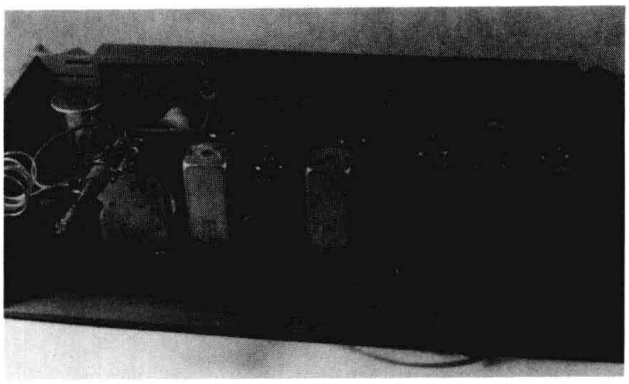


Figure 3.5 An AC/DC Radio Receiver of the 1950s Tube filaments (or heaters) are connected in series across the 110 volt line.

OHM'S LAW

Ohm's Law: The current in amperes that will flow in any circuit is directly proportional to the pressure applied in volts and inversely proportional to the circuit resistance in ohms.

Perhaps the most important law of electricity and electronics, and the one most widely used, was first stated by George Ohm in 1827. Ohm pointed out the simple but important relationships that exist between electrical potential, electrical resistance, and the strength of the current flowing. Ohm's Law was important because, once this mathematical relationship was spelled out, other new laws could be developed—for example, Kirchhoff's voltage law and Norton's and Thevenin's statements of current and voltage in complex resistive circuits.

Ohm conducted his research with rather primitive laboratory equipment. Fundamental in his studies were lengths of wire that he had drawn through dies with his own hands. He soon discovered that resistance is basically a function of (a) the length of the wire used, (b) the wire's cross section, and (c) the material from which the wire was drawn.

For example, if in a certain circuit the resistance is held constant and the voltage is varied, doubling the voltage will produce a doubling of the current flowing, while halving the voltage will halve the current. Holding the voltage constant and varying the circuit resistance has exactly the opposite effect: Doubling the resistance will halve the current; halving the resistance will double the current flow.

Three Basic Statements of Ohm's Law

The first basic statement of Ohm's Law is usually reduced to "electrical shorthand" to read as follows:

$$I = E/R$$

In this statement, I stands for current *intensity* in amperes, E stands for *electromotive force* in volts, and R stands for the circuit *resistance* in ohms. This formula holds true only if the current is in *amperes*, the electrical pressure is in *volts*, and the resistance is in *ohms*. If, for example, the current was stated in milliamperes, the voltage in kilovolts, or the resistance in megohms, it would not be possible to substitute and solve any of the three statements of Ohm's Law until the quantities were converted to amperes, volts, and ohms.

The other two statements of Ohm's Law are obtained by simple algebraic rearrangement of the original statement. They are as follows:

$$E = IR$$

$$R = E/I$$

Once you have memorized these three basic statements, you are able to perform three vital operations in electronics: finding the *value of current* that will flow if the voltage and resistance are known, finding the *circuit voltage* when the current and resistance

are known, and finding the *circuit resistance* when the voltage and current are known.

Problems and Solutions

The problems that follow show how Ohm's three formulas can be put to work. The examples are typical in terms of the values of voltage, current, and resistance encountered. Remember that you always have to know *two* of these quantities before you can find the third.

EXAMPLE 1
The filament of a sealed-beam headlamp used on a modern car has a resistance of 2 ohms. How much current in amperes will it draw when connected across a 12 volt automotive battery?

Solution
Select the proper equation:

$$I = E/R$$

Substitute and solve:

$$I = 12/2 = 6 \text{ amperes}$$

EXAMPLE 2
The heater coil of a certain vacuum tube has a resistance of 21 ohms. When connected across an unknown voltage source, it draws exactly 0.3 amperes. What is the voltage connected across its terminals?

Solution
Select the proper equation:

$$E = IR$$

Substitute and solve:

$$E = 0.3 \times 21 = 6.3 \text{ volts}$$

EXAMPLE 3
A soldering iron, when connected to the 110 volt line, draws a current of 0.5 amperes. What is the resistance of its heating element?

Solution
Select the proper equation:

$$R = E/I$$

Substitute and solve:

$$R = 110/0.5 = 220 \text{ ohms}$$

A common situation found in electronic circuits is that currents are frequently too small to be expressed readily in amperes and are expressed instead in *milliamperes* (abbreviated either mils or mA). When current is stated in mils, it must, of course, be converted to amperes before the problem can be solved. Since a mil is one-thousandth of an ampere, to convert mils to amperes you divide the current stated by 1000—that is, move the decimal point three places to the left.

EXAMPLE
The current flowing through a certain resistor is measured as 40 mils. The resistor is color coded 270 ohms. What is the voltage across the ends of the resistor?

Solution
Convert 40 mils to amperes:

$$40/1000 = 0.04 \text{ amperes}$$

Select the right formula:

$$E = IR$$

Substitute and solve:

$$E = 0.04 \times 270 = 10.8 \text{ volts}$$

After we have studied alternating current (Chapter 5), we will show how Ohm's Law can apply to AC as well as DC circuits. When applied to AC circuits, the symbol Z (impedance) is substituted for the symbol R (resistance).

POWER COMPUTATIONS

There are no specific laws for electrical power. However, the measurement and calculation of power are closely related to Ohm's Law.

Power and Wattage

Power is the rate at which energy is expended, or used up. Expending a great deal of electrical energy in a very short period of time represents the expenditure of a very large amount of power. The same amount of energy spread over several hours would represent a very modest expenditure of power.

A **watt,** named after James Stephenson Watt (1736–1819), inventor of the steam engine is the amount of electrical energy expended when 1 volt drives a current of 1 ampere's strength through a circuit having a resistance of 1 ohm. In certain applications—for example, determining how much heat is generated in a certain resistor by the passage of an electrical current—the basic formula is tied to Ohm's Law:

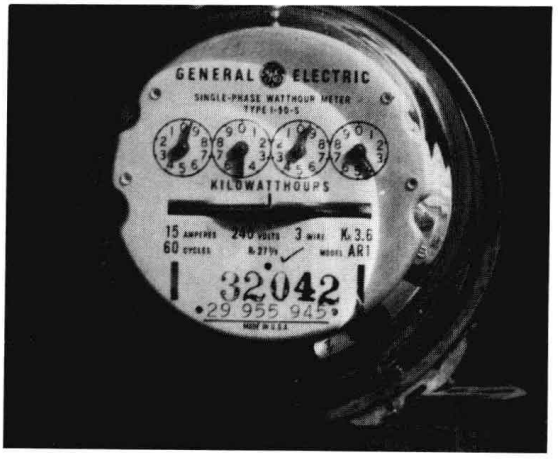
$$\text{power} = E \times I$$

This equation will yield a product that represents the instantaneous amount of power generated within the resistor by the passage of the current. This power figure must always be *less* than the power rating of the resistor—say, 7 or 8 watts for a 10 watt-rated resistor—or the resistor will overheat and eventually burn up.

The Time Factor

Computing the amount of electrical power consumed by a radio transmitter, an electrical motor, or a residence or business always involves use of a *time frame*. The basic unit is always the consumption of 1 watt for the period of 1 hour and is known as the **watt-hour.** Since a watt-hour is too small a yardstick for measuring most electrical consumption, the **kilowatt-hour** is used instead—that is, the consumption of 1000 watts of power continuously for exactly 1 hour. For example, a microwave oven will consume about 1.5 kilowatts, while a clothes dryer may consume as much as 5.0 kilowatts (usually stated as 1.5 kW and 5.0 kW). In the Pacific Northwest, where abundant hydroelectric power is relatively inexpensive, the cost per kilowatt-hour might be as low as three cents. It is much more expensive elsewhere.

When we discuss the power output of nuclear and hydroelectric power plants, we find that even the kilowatt is too small a unit of measure, so we use the **megawatt** (MW), which equals 1,000,000 watts. For example, the Trojan Nuclear Plant at St. Helens, Oregon, considered one of the nation's largest, normally produces about 1200 megawatts of electrical power. Bonneville Dam, which spans the Columbia River just east of Portland, produces about 1000 megawatts.



A GE wattmeter is actually a form of electrical motor. The wattmeter drives the individual indicating hands through a series of reduction gears. Like many modern meters, its readout is in digital form.

Three Basic Formulas for Computing Power

Only three basic formulas are necessary to compute the power in any DC electrical circuit. They are as follows:

$$P = EI$$

$$P = I^2R$$

$$P = E^2/R$$

For these equations to hold true, the electrical pressure must be expressed in *volts*, the current strength in *amperes*, and the resistance in *ohms*. If they are stated in other units, they must be converted to volts, amperes, and ohms before being substituted in the formula. This formula is especially useful for approximating the amount of current drawn from the power line by medical equipment, household appliances, and even radio transmitters.

By rearranging the original statement, $P = EI$, we obtain two other helpful formulas. These formulas have many useful applications in determining the voltage or amperage drawn by a piece of electronic equipment if its power rating and one of the other two quantities are known. The additional formulas are as follows:

$$I = P/E$$

$$E = P/I$$

Some simple problems follow.

EXAMPLE 1
A 100 watt soldering iron is plugged into a 110 volt outlet. How much current will it draw in amperes?

Solution
Select the right formula:

$$I = P/E$$

Substitute and solve:

$$I = 100/110 = 0.9 \text{ amperes}$$

EXAMPLE 2
An incandescent lamp carrying a rating of 300 watts draws a measured 1.36 amperes from the power source, which is unknown. What is the voltage of the power line across which the bulb is connected?

Solution
Select the right formula:

$$E = P/I$$

Substitute and solve:

$$E = 300/1.36 = 220 \text{ volts}$$

ELECTRICAL CIRCUITS

Three basic circuits are used in electronics: series, parallel, and series-parallel. No matter how complex the circuitry used, it can always be broken down into these three primary circuits or into combinations of them. In addition to the three basic circuits, two other types of circuits are common: open circuits and short circuits.

Open and Short Circuits

Figure 3.6 is an example of an **open circuit.** A power source is connected to an electric lamp with a burned-out filament. Since electrons can't flow across the gap in the filament, there is no current flow, and we have an open circuit. The circuit can be fixed by replacing the burned-out bulb with a good one. Burned-out fuses also cause open circuits, but they normally are not replaced until the reason they "blew" has been determined.

In Figure 3.7, we have replaced the defective lamp with a good one, but this time we have placed

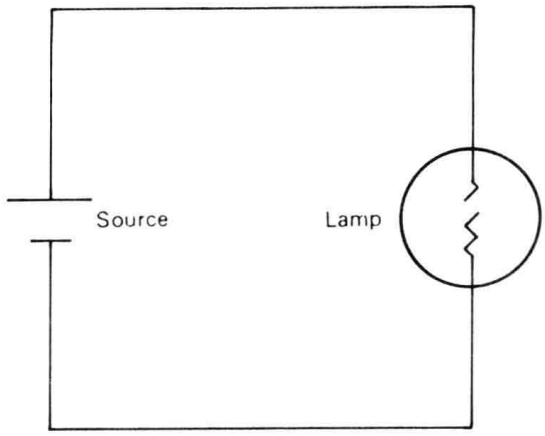


Figure 3.6 Open Circuit The lamp filament is burned out.

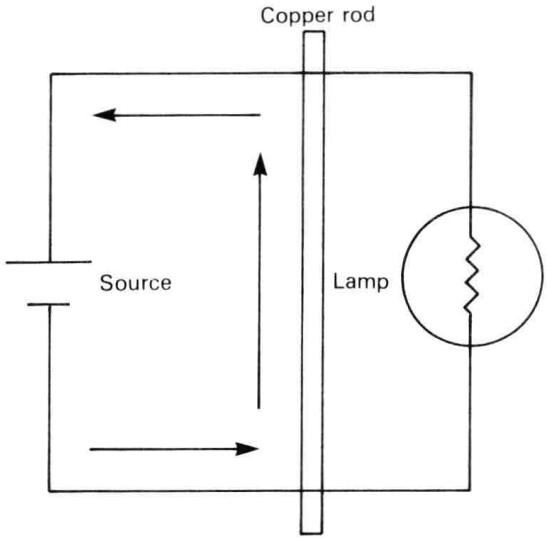


Figure 3.7 Short Circuit A copper bar placed across the lamp terminals produces a short circuit, and the lamp does not light.

a heavy copper rod directly across the lamp terminals. Once more we find that the lamp fails to light, but this time for a different reason—a **short circuit.** Electrons always take the path of least resistance. Therefore, the current will flow through both the power line and the copper bar as the line current takes the short (hence "short circuit"), or direct, path from one conductor to the other without flowing through the load (the lamp). The tremendous flow of current will probably cause the line fuses to blow or the circuit breakers to trip.

Fuses and **circuit breakers** are the safety valves of an electrical circuit. Fuses are made of fusible links of a metal of comparatively high resistance and low melting point. When the current becomes excessive, the fusible link melts to open and protect the circuit. Circuit breakers work on electromagnetic principles. When the current through the turns of wire of the circuit breaker's electromagnet exceeds a safe value, its "pull" opens a switch that, in turn, opens the circuit. Both fuses and breakers are connected in series with the circuit they are protecting.

Series Circuits

Series circuit: A circuit that has a single path over which current flow can take place.

Series circuits are the simplest of circuits. They are less common today than in the 1940s. Many radio receivers of that period used just five vacuum tubes with their heaters connected in series across the 110 volt power line. Series circuits did not improve a set's operation but did greatly reduce cost by eliminating the need for a power transformer. Christmas tree lights of the same period used ten incandescent lamps with 12 volt filaments in series across the 110 volt line. Even the heavy-duty lamps used in street lighting were frequently connected in series across a 500 volt source. The problem with such circuits was that if just one unit failed, the entire system became inoperative, and the defective unit had to be replaced to restore operation.

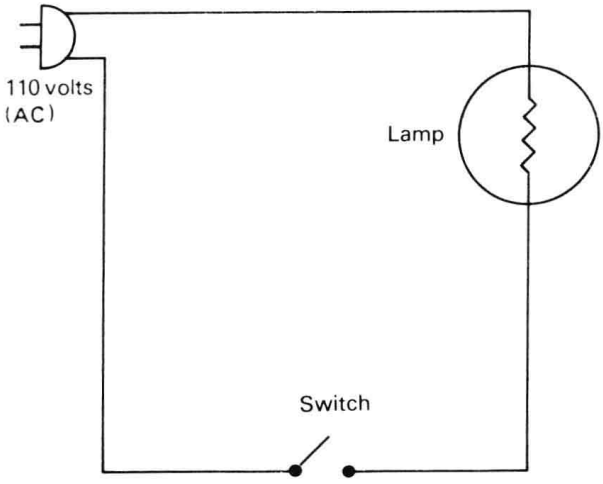


Figure 3.8 Off-On Switch on a Lamp A switch is always in series with the device it controls.

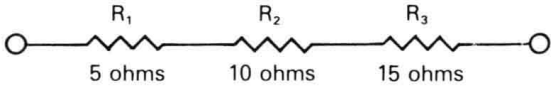


Figure 3.9 Series Circuit The total resistance is the same as that produced by a single 30 ohm resistor.

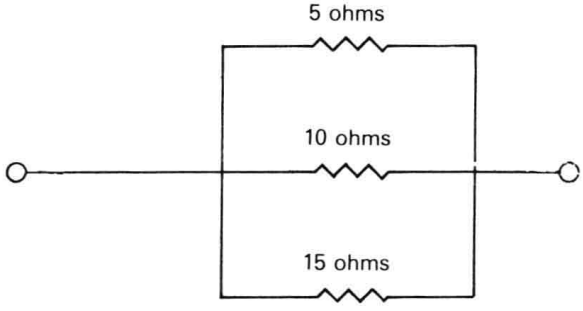


Figure 3.10 Parallel Circuit Here the current has three paths over which it can flow.

Today, virtually the only application of the series circuit is to provide control. The off-on switch on a TV set, a stereo, or a lamp (Figure 3.8); the pop-up switch on an electric toaster; and the ignition switch on a car are always connected in series with the circuit being controlled. When the switch is closed, the equipment operates; when it is open, everything stops. There is one further application of series circuits: Devices designed to protect electronic equipment against overload, such as fuses and circuit breakers, are always connected in series with one side of the line leading to the power source.

EXAMPLE

In Figure 3.9, a 5 ohm, a 10 ohm, and a 15 ohm resistor are connected in series. Find their effective total resistance in the circuit.

Solution

Start with the formula for adding the resistors:

$$R_T = R_1 + R_2 + R_3$$

Substitute and solve:

$$R_T = 5 + 10 + 15 = 30 \text{ ohms}$$

The total is the same as is produced by a single 30 ohm resistor.

Parallel Circuits

Parallel circuit: A circuit that has more than one path over which current flow can take place.

A parallel circuit is shown in Figure 3.10. Note that there are three paths over which current flow can

take place. Most current will flow through the 5 ohm resistor, the next largest amount through the 10 ohm resistor, and the least current through the 15 ohm resistor. However, since some current flows through each path, the circuit resistance will be *less* than that of the lowest value resistor.

Parallel circuits have wide application in the home. Appliances such as floor lamps, food mixers, space heaters, vacuum cleaners, and television receivers plug into receptacles connected to parallel circuits made up of number 12 gauge (or larger) copper wire. The wire brings 110 volt power at some limited amperage from the **entrance box,** a metal or plastic box where the electric power enters the building from public utility lines. There may be eight to ten parallel circuits in the entrance box to provide power for the usual outlet circuits plus high-current power for appliances such as electric stoves and clothes dryers. Normally the outlet circuits are protected by circuit breakers, magnetic switches that open when too much current is drawn by a circuit either because of a "short" or because too many appliances have been plugged into the circuit. The higher-current appliances, such as stoves and clothes dryers, are protected by cartridge-type fuses.

Parallel circuits are widely used in industry and business. Heavy-duty motors are always fed from their power source over a parallel circuit and are provided with their own circuit breakers and off-on switches. Fluorescent lighting, in wide use in industry and in business buildings, is invariably powered by a parallel circuit, but with the off-on switch in series with one conductor in the circuit.

Parallel circuits have one other application in industry. Electronic technicians, working under the supervision of a design engineer, usually wire and assemble the preliminary model of a new electronic device, or **prototype.** Often they need a particular value of resistor, or resistance, unavailable in the stockroom. The solution is to start experimenting with various values of resistors in parallel until they find just the right resistance value. If this approach doesn't work out well with the resistor stock in hand, technicians may try a dual approach: connecting resistors in *parallel* and then experimenting with various values of resistors to connect in *series* with the parallel circuit. It isn't unusual in a prototype model to find three or four resistors connected to provide the exact value of resistance that will be provided by a single resistor in the final production model.

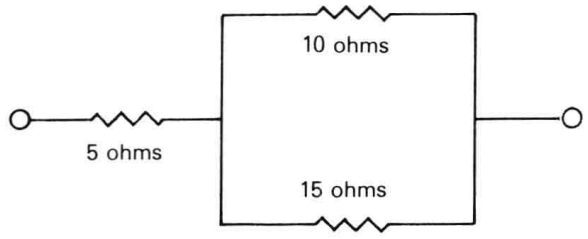


Figure 3.11 Series-Parallel Circuit The 5 ohm resistor is in series with a 10 and a 15 ohm resistor in parallel.

Series-Parallel Circuits

Series-parallel circuit: A circuit that is a combination of both series and parallel circuits.

The circuit shown in Figure 3.11 is typical: a 5 ohm resistor in series with a 10 ohm and a 15 ohm resistor in parallel. Finding the effective resistance value of the combination is easy. Total circuit resistance of such a combination will always be at least 5 ohms plus the added resistance of the two in parallel.

EXAMPLE

Find the circuit resistance of a 5 ohm resistor in series with a 10 and a 15 ohm resistor connected in parallel.

Solution
Find the effective resistance of the two parallel resistors.

$$R_T = (R_1 \times R_2)/(R_1 + R_2)$$

Substitute and solve:

$$R_T = (10 \times 15)/(10 + 15) = 150/25$$
$$= 6 \text{ ohms}$$

Add the value of the series resistor:

$$6 + 5 = 11 \text{ ohms}$$

The circuit would work identically with a single 11 ohm resistor in series with the circuit.

RESISTANCE AND RESISTORS

Resistance: The opposition offered by a circuit to the flow of electrons through it.

All conductors, even silver, have some resistance. When the conductor is in the form of wire, its resistance becomes a function of both the length of the wire and the area of a cross section. For example, if 100 feet of copper wire of a certain diameter has a resistance of 1 ohm (a reasonable figure), 500 feet will have a resistance of 5 ohms, and 50 feet will have a resistance of 0.5 ohm. If we substitute a piece of wire of identical length but with twice the diameter of the first piece, we find that the resistance has dropped to 0.25 ohm.

In other words, the resistance of a conductor *decreases* as the cross-sectional area of the wire (determined by the formula $A = \pi r^2$) *increases*. For now, we can assume that electrical current flow takes place uniformly throughout the cross section of the conductor. (This assumption must be modified in the discussion of receiver and transmitter circuits in Chapter 7. At radio frequencies, currents concentrate on the surface to produce what is known as "skin effect.")

Resistive Materials

Resistor: An electronic component made of a material that offers resistance to electron flow.

When we discuss substances characterized by resistance, we are usually talking about materials such as manganin, nichrome, and carbon, which are used to manufacture devices called *resistors*. Substances that exhibit resistance have their atoms so arranged that, as electrons move from the negative to the positive pole of the circuit, many collisions between the atoms and the moving electrons take place.

Each collision turns some of the electrical energy into heat energy and in effect reduces the amount of electrical energy available. Resistance is not the same as the quality of *nonconductance*. Substances such as paper and wood do not conduct an electrical current readily because they have no appreciable numbers of the free electrons essential for current flow. Resistive materials do provide free electrons, but it is the frequent collisions between their atoms and the major flow of electrons from the negative to the positive pole of the circuit that produce the opposition to electrical current flow.

When we talk about resistors in electronics, we're talking about electrical components, or parts, in which a specific amount of resistance is packaged in each separate unit. Resistance is always stated in ohms, and the exact resistance is commonly indicated by a color code. The quality of resistance is always relative since all conducting materials—even the very

best conductors (silver, copper, gold, and aluminum)—have *some* resistance.

Carbon Resistors

In electronics, the most commonly used substance in the making of resistors has been the element carbon. Carbon is a fairly good conductor of electricity. However, if a thin film of carbon is deposited on a glass rod, a very stable resistor results. Its exact resistance will depend on the thickness of the deposited coating and on the length and the circumference of the glass rod. Such a resistor is too expensive to be used freely in most applications.

Carbon resistors can be made much more cheaply by mixing various proportions of powdered carbon with bakelite powder and then subjecting the mixture to high pressure and heat. If the ratio of carbon to bakelite powder is high, a low value of resistance results. A large ratio of bakelite powder to carbon produces a high value of resistance. By carefully controlling the ratio of the two principal ingredients, the right ohmmic value for the resistor can be produced at extremely low cost. A resistor unit made in this way is always referred to as a **carbon resistor.** In large-quantity purchases, the unit cost is frequently just a few cents.

Most of the resistors you see under the chassis of a TV set or a stereo amplifier are carbon resistors. While most carbon resistors are quite small, some may be quite large. Size has to do with **wattage rating**—that is, the resistor's ability to radiate into the surrounding air the heat generated within by the current flowing. The larger the resistor, the more heat it can expel and the greater its rating in watts. Where the amount of heat generated is insignificant, values of 1/2 watt or less are widely used.

Carbon resistors are made with wattage ratings up to 2 or 3 watts. They may have the diameter of a pencil and be several inches long. Since the resistance of carbon tends to *decrease* as its temperature *increases*, the hotter it gets, the greater the current flow becomes. Eventually the resistor overheats and may actually burn in two. Carbon resistors are therefore not used in circuits where a heavy current is to be carried.

Wire-Wound Resistors

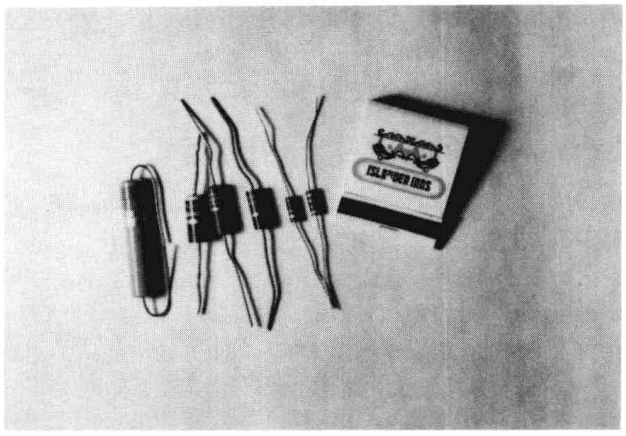
In circuits where heavy currents must be handled, **wire-wound resistors** are almost invariably used. For example, the burners on an electric stove, the heating unit in a hair dryer, and the heating element in a soldering pencil all contain wire-wound resistors. The wire used is almost always nichrome, a patented **alloy** (a mixture of two or more metals) made of nickel and chromium.

The resistance of a wire-wound resistor made of nichrome is adjusted by controlling the diameter and the length of the resistance wire. High resistance is produced by a long length of fine nichrome wire, low resistance by a short length of large-diameter nichrome wire. The wire is usually supported on a mica or a ceramic base. A ceramic base may be covered with a ceramic glaze and then fired in a furnace. Resistance values ranging from 1 to 100,000 ohms or higher are readily available from stock in wattage ratings from 5 watts to 100 watts and higher.

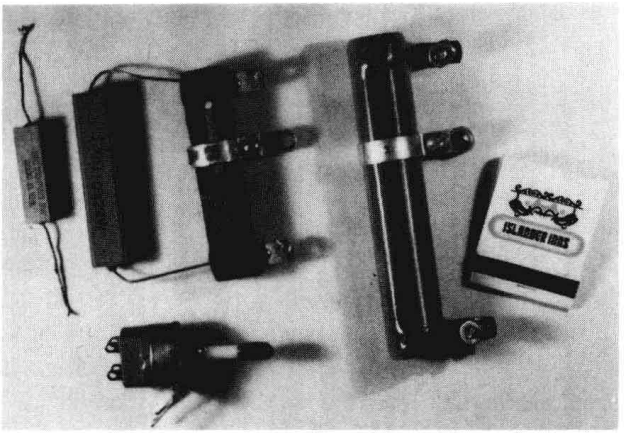
Wire-wound resistors can be operated for hundreds of hours without failure at temperatures above 100°C. They are much more expensive than carbon resistors and therefore are used only in comparatively high-current circuits where absolute reliability is of maximum importance. Figure 3.12 shows some typical carbon and wire-wound resistors.

Variable Resistors

In the early days of radio broadcasting, **variable resistors** were used principally to control the filament voltage of both receiving and transmitting tubes. They were usually made of a C-shaped section of bakelite, over which nichrome wire had been wrapped, and a wiping arm in contact with the resistance element. As the shaft attached to the wiping arm was turned, more or less resistance appeared between one end of the resistance winding and the connection made to the moving arm. This changing resistance



A. Carbon resistors ranging from 4 watts (left) to 1/2 watt (right).



B. Wire-wound resistors ranging from 5 watts (left) to 20 watts (right). Larger units are semiadjustable. In the foreground is a potentiometer, or variable resistor, of the carbon type.

Figure 3.12 Typical Carbon and Wire-Wound Resistors

could be used to apply exactly 5 volts to the delicate filament of the receiving tubes that usually were powered by a 6 volt storage battery.

Variable resistors have wide application in many modern electronic products. For example, the volume control on transistor radios, television sets, and stereo amplifiers, as well as the "gain control" on a CB or an amateur radio transceiver are all variable resistors. On a TV set the controls labeled *horizontal hold, vertical hold, vertical linearity,* and *brightness,* are actually variable resistances of various values.

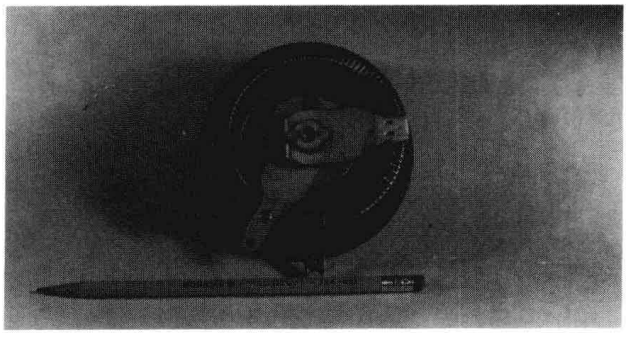


Figure 3.13 A Large Wire-Wound Variable Resistor Used to Control the Filament Voltage of Transmitting Tubes Note how the wiping arm makes contact with the resistance strip.

Variable resistors are commonly made by spraying a controlled thickness of carbon film on a C-shaped bakelite strip. A connection is made to each end of the strip, and a third connection is made to a wiping arm that moves across the strip when its shaft is turned. Wire-wound types are also available for high-current circuits, but they are much more expensive. Note that the resistance between the outside ends of the resistance strip is always constant, while the resistance between either outside lug of the resistance strip and the moving arm varies as the shaft is turned. The usual terminal arrangement is to use the two outside lugs to connect to the resistance strip and the center lug to connect to the wiping arm (Figure 3.13).

A variable resistor with three terminal lugs brought out for circuit connection is technically called

a **potentiometer** but commonly referred to as a *pot* or *control*. Variable resistors have a number of applications—as dimmers to control the intensity of area lighting, as controls to vary the rate of cooking in a microwave oven, and as sensitivity controls for proper adjustment of electronic garage-door openers.

Controls that have become noisy or intermittent in operation can frequently be restored to proper operation through the use of an aerosol spray intended for the cleaning of electrical contacts. Spray both the shaft bearing and the resistance strip, then rapidly rotate the shaft before the cleaner-lubricant spray has time to evaporate. Normally it is not necessary to disconnect any wiring to the control's terminals for cleaning and lubrication.

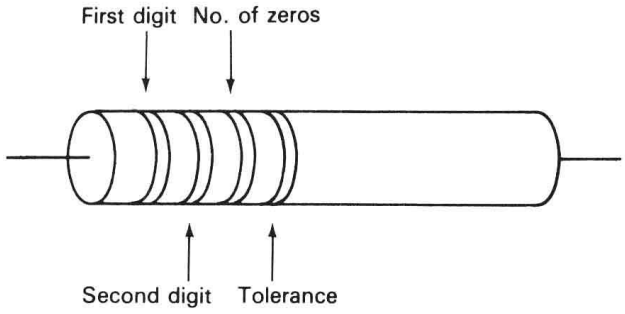


Figure 3.14 Color Bands on a Resistor Always read the color code from left to right, starting at the end to which the bands are closer.

Color Coding Resistors

A necessary skill for all electronic technicians, and a desirable one for anyone with an avocational interest in electronics, is to identify quickly the value in ohms of a certain resistor by interpreting its color code. Color codes for resistors have been in existence for at least forty years, but they were standardized during World War II and have remained substantially without change since that time.

The Electronic Industries Association (EIA) color coding system is based on a table of the following ten colors and their number equivalents:

Black 0
Brown 1
Red 2
Orange 3
Yellow 4
Green 5
Blue 6
Violet 7
Gray 8
White 9

Referring to Figure 3.14 (a typical carbon resistor), note that there is a series of color bands starting closer to one end of the resistor than to the other. A resistor usually has four bands. The color code is always read from the end to which the color bands are closest. Following the sequence in Figure 3.14, the first stripe is the first digit of our resistor's value. The second stripe is the second digit, while the third stripe tells how many zeros follow these two digits. The fourth stripe, always a metallic color (silver or gold), indicates the tolerance, or accuracy, of the resistor and does *not* enter into the calculation of its resistance.

EXAMPLE 1
Find the resistance in ohms of a resistor coded yellow-violet-black-silver.

Solution
Yellow is 4, so the first digit is 4. Violet is 7, so 7 is written after the 4. Black stands for 0, so no zeros follow. The value of the resistor, then, is 47 ohms. Silver indicates that accuracy is guaranteed to be within $\pm 10\%$ of the value coded. The answer is 47 ohms $\pm 10\%$.

EXAMPLE 2
Find the resistance in ohms of a resistor coded red-red-red-gold.

Solution
Red is 2, so both the first and the second digits are 2. The third red, which is also 2, indicates that two zeros follow. The value of the resistor, then, is 2200 ohms. The gold stripe indicates that accuracy is guaranteed to be within ±5% of the value coded. The answer is 2200 ohms ±5%.

EXAMPLE 3
Find the resistance in ohms of a resistor coded brown-black-green-silver.

Solution
Brown is 1, so the first digit is 1. Black is 0, so 0 is written after the 1. Green is 5, so five zeros follow. Silver indicates that accuracy is guaranteed to be within ±10% of the value coded. The answer is 1,000,000 ohms ±10%.

While it is technically correct to write the answers in ohms (as in the preceding examples), different practices apply on a wiring diagram, or **schematic.** The usual practice for values under 1000 ohms is to write down the number and follow it with the Greek letter omega. For values of 1000 ohms to 1,000,000 ohms, the practice is to divide by 1000 and write the number followed by a lowercase k for thousand. For example, 2200 is written 2.2k. Values of 1,000,000 or more ohms are expressed in **megohms** (1,000,000 ohms), which can be spelled out in full or abbreviated M ohms.

When a resistor is valued at less than 10 ohms, it is marked with a modified color code. For example, suppose you have a resistor color-coded from left to right as follows: brown-red-gold-silver. First write down 1 and 2, the first two digits. The third stripe, gold, is a "multiplier" with a value of 0.1. Multiplying 12 by 0.1 yields 1.2 ohms, the coded value. The fourth stripe, silver, shows that the tolerance is ±10% of the coded value. Were the third stripe to be silver instead of gold, you would multiply the value of the first two digits by 0.01.

No relationship exists between the size of a resistor and its resistance in ohms. A 1/4 watt resistor may measure 10 million ohms. However, a definite correlation exists between physical size and wattage rating, which is linked to the resistor's ability to get rid of the heat generated by current flow. In general, the larger the resistor, the greater the wattage rating.

SUMMARY

Electrical activity can be described as the effort of unlike electrical charges to move toward each other and the effort of like charges to move away from each other. In the words of Coulomb's Law of Charges: Unlike charges attract, and like charges repel; the amount of attraction or repulsion is a function not only of the size of the charges but also of the spacing between them.

Kirchhoff's Law of Current says that no more (or less) current can leave a junction of conductors than enters the same junction. In analyzing a specific circuit, an electronic technician can always assume that the current entering a junction is identical to the current leaving that junction. Kirchhoff's Law of Voltage says that if a source is connected to a load, such as a number of resistors connected in series, all the voltage provided by the source will be used up in driving the electron flow (current) through the resistors. As the electrons encounter the resistive material, collisions take place that turn part of the voltage energy into heat energy and subtract from the amount of voltage energy available. Kirchhoff referred to this loss of voltage across each individual resistor as a *voltage drop* and stated that the sum of all the individual voltage drops around a circuit is exactly equal to the voltage of the source.

Ohm's Law, the most important statement in the field of electronics and electricity, says that the current flow (in amperes) in any circuit is directly proportional to the applied pressure (in volts) and inversely proportional to the resistance encountered (in ohms). It holds true only when the pressure is expressed in volts, the current in amperes, and the

resistance in ohms. When these values are stated in other units, the units must be converted to volts, amperes, and ohms before the problem can be solved.

Ohm's Law is also related to the determination of electrical power: When 1 volt drives 1 ampere through a circuit having a resistance of 1 ohm, 1 watt of electrical power is expended. If this activity took place continuously for an hour, 1 watt-hour would be consumed. Since a watt-hour is a very small amount of energy, the usual unit of measure is the kilowatt-hour, in which 1000 watts of electrical energy are expended for an hour.

All electronic circuits can be assigned to one of three categories: series, parallel, or series-parallel. No one circuit is best; the circuit that should be used is the one that does the proper job for a particular application. A series circuit has a single path over which current flow takes place; a parallel circuit has two or more paths over which current flow takes place; and a series-parallel circuit is a combination of series and parallel circuits.

Resistors are the most frequently encountered component in an electronic circuit or a piece of electronic equipment. They offer a specific opposition to the flow of an electrical current stated in ohms. Fixed resistors may be of either the carbon or the wire-wound type. Their wattage rating expresses how much electrical energy they can convert into heat energy without excessive temperature rise. Variable resistors, usually called potentiometers, have wide application in electronics. They are used as the volume and tone controls on stereo amplifiers and as brightness and color-balance controls on television sets.

KEY WORDS AND CONCEPTS

alloy
carbon resistor
circuit breaker
Coulomb's Law of Charges
entrance box
fuse
kilowatt-hour
Kirchhoff's Law of Current
Kirchhoff's Law of Voltage
megawatt
megohm
Ohm's Law
open circuit
parallel circuit
potentiometer
power
prototype
resistance
resistor
schematic
series circuit
series-parallel circuit
short circuit
variable resistor
voltage drop
watt
watt-hour
wattage rating
wire-wound resistor

SELF-TEST QUESTIONS

1. In your own words, state Coulomb's Law of Charges.
2. In your own words, state Kirchhoff's Law of Current.
3. As simply as possible, describe what is meant by Kirchhoff's loop law or voltage law.
4. Using plain language, not formulas, give the basic statement of Ohm's Law.

5. In the Ohm's Law equation $I = E/R$, what does each letter stand for?
6. In analyzing a circuit through the use of Ohm's Law, you find that electrical pressure, resistance, and current flow are expressed in quantities other than volts, ohms, and amperes. What step must you take before proceeding further?
7. In your monthly bill from your electric power utility, upon what electrical unit is the charge computed?
8. Define a series circuit.
9. Define a parallel circuit.
10. Define a series-parallel circuit.
11. Is the off-on switch on a TV set connected in series or in parallel with the 110 volt power circuit?
12. Why are carbon resistors so widely used in all types of electronic equipment when wire-wound resistors can carry heavier currents and offer more protection against overloads?

4 Basic Electro-magnetism and DC Theory

This Shure dynamic microphone is in use in a radio control room. The microphone operates essentially in reverse of the action of the permanent magnet loudspeaker.

Magnetism
- Magnetic Poles
- Magnetic Materials
- Temporary and Permanent Magnets
- Magnetic Field Behavior
- Theories of Magnetism
- Magnets as Complete Systems
- Types of Permanent Magnets

Electromagnetism
- Strength of an Electromagnet
- Applications of Electromagnetism

Inductance
- How Inductance Works
- Inductive Reactance
- Self-Inductance and Mutual Inductance
- Solenoids
- Measuring Inductance

Capacitance
- Factors That Determine Capacitance
- Variable Capacitors
- Charging and Discharging Capacitors
- Units of Capacitance
- Naming Capacitors
- Electrolytic Capacitors

Voltage Rating of Capacitors
Capacitors in Series and in Parallel
Precautions in Connecting Capacitors
Capacitive Reactance

Combining Inductors and Capacitors
Tuning and Tuned Circuits
Calculating Inductance and Capacitance

Summary

Key Words and Concepts

Self-Test Questions

While the topics in the chapter title might at first glance seem quite unrelated, in fact they are closely related. For example, it is impossible to distinguish between the magnetic field around a permanent magnet and the magnetic field caused by the passage of current through a coil (electromagnetism). Also, to understand the behavior of a coil (inductance), we must first be familiar with the effects on the coil of the changing magnetic field around it. Capacitors and capacitance are related to a discussion of inductance since coils (called *inductors*) and capacitors are most frequently used together in electronic circuits to produce both filtering and tuning. Inductances and capacitances have another similarity: They store energy in the form of electromagnetic and electrostatic fields.

Filtering: Separating voltages or currents of two or more differing frequencies from each other.

Tuning: The process of selecting the desired station and rejecting undesired stations.

MAGNETISM

Our knowledge of magnets extends back over thousands of years. Naturally occurring magnets were then rather commonly found on the surface of the earth, particularly in an area of Asia Minor known as Magnesia, where iron-bearing sediments had been laid down in magnetic patterns under the influence of the earth's magnetic field. The name of this province gave us such words as *magnet, magnetism,* and *magnesium* (a silvery white metal widely used in aircraft fabrication and in medicine).

Much later, when the lakes dried up and the sediments hardened into rock, these materials still maintained their magnetic quality. A lump picked up from the ground, for example, could attract a nearby piece of iron. Suspended by a thread, a lump of magnetic material would tend to line up with the earth's ever-present magnetic field and point toward the earth's north and south geographic poles. Therefore, these substances were called lodestones, or leading-stones. Outside of their attraction for iron and their ability to indicate the earth's north and south magnetic poles, lodestones were of little use since they were both weakly and unevenly magnetized.

The earth appears to be unique among the planets of our solar system in having a large, strongly magnetized iron core at its center. The magnetic field that this mass produces extends tens of thousands of miles into space beyond the earth's surface (Figure 4.1). The most important effect of this magnetic field is that it harmlessly deflects into outer space electrical particles radiated from the sun. Were these particles to reach the earth's surface, they could cause great damage to all forms of life.

Magnetic Poles

Strangely, the earth's south magnetic pole is near the earth's north geographic pole, while the north magnetic pole is near the south geographic pole. We already know that unlike magnetic poles attract and

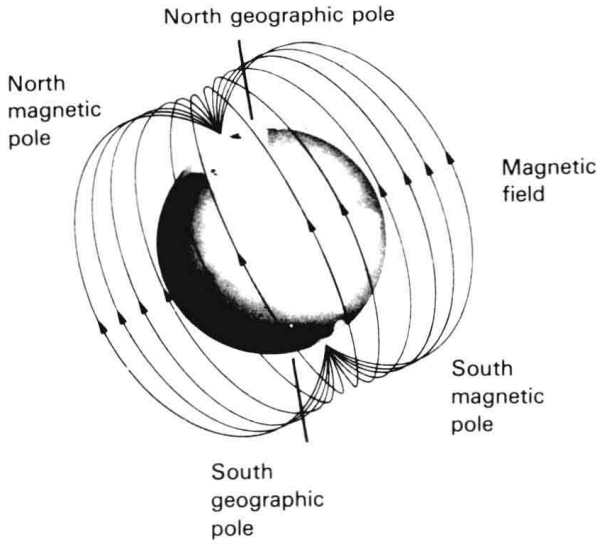


Figure 4.1 Earth's Magnetic Field The earth, because of its intensely magnetized iron core, produces an extensive magnetic field as it spins on its axis.

that the pointer on a magnetic compass is the north pole of a pivoted permanent magnet. Therefore, when the compass points north, it is because the pointer is being attracted by the earth's south magnetic pole. (The magnetic poles, in practice, are named for their adjacent geographic poles.) Both magnetic poles are several thousand miles from the geographic poles, so only at certain locations on the earth's surface does the compass needle point exactly north and south. This discrepancy has to be taken into consideration when a compass is used for navigation.

Magnetic Materials

Experimenters have now known for well over a century that only three elements—iron, nickel, and cobalt—have pronounced magnetic properties. That is, only these three are strongly attracted or repelled by a magnetic field, and only these three or their alloys can be used in the manufacture of permanent magnets. An **alloy** is a mixture of two or more metals usually melted together. An alloy is *not* a chemical compound.

Temporary and Permanent Magnets

Temporary magnet: An object that fails to retain its magnetic properties after the magnetizing force has been withdrawn.

Permanent magnet: An object that retains its magnetic properties over a long period of time.

Magnetic effects may be either temporary or permanent. For example, when we pick up a soft iron nail with a bar magnet, the nail itself temporarily becomes a complete magnetic system with both a north and a south magnetic pole. As such, it can pick up other nails or objects made of the three magnetic elements. But when the magnetizing force (the bar magnet) is withdrawn, within just a few seconds the nail loses nearly 100% of its magnetic qualities. Soft, pure iron is the best example of a material that is readily magnetized but fails to retain its magnetism after the magnetizing force has been withdrawn.

If we bring a bar magnet in contact with the high-carbon steel blade of an ordinary screwdriver, the screwdriver becomes a complete magnetic system with both a north and a south pole. Like a soft iron nail, it will attract objects made of magnetic materials, but with one important difference. Take the bar magnet away, and the steel blade still retains its magnetic characteristics. The ability to retain magnetic properties after the magnetizing force has been withdrawn is referred to as **retentivity.**

The first permanent magnets were made of steel, which is basically iron to which trace amounts of

carbon and certain metals have been added. In more recent years aluminum-nickel-cobalt and cobalt with barium carbonate added have been widely used in the manufacture of powerful permanent magnets.

Magnetic Field Behavior

Magnetic field: The area where the magnetic effects of an object are readily observable.

The behavior of either a temporary or a permanent magnet can best be explained in terms of the magnetic field associated with it. Referring to the illustrations of the magnetic fields surrounding a bar magnet and a horseshoe magnet (Figures 4.2 and 4.3), we need to understand that what we see is only an outline, not the field itself. Each one of the tiny iron filings behaves like a temporary magnet and turns itself so that the **lines of force** that comprise the field flow through its miniaturized magnetic system. In arranging themselves in this manner, the filings also outline the magnetic field that surrounds the permanent magnet. The filings indicate the field's scope as well as its areas of greatest and least intensity.

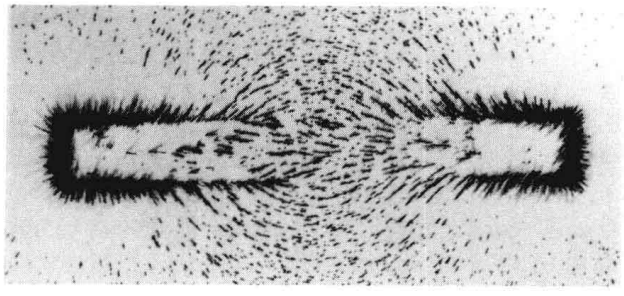


Figure 4.2 Bar Magnet Iron filings, acting as temporary magnets, concentrate along lines of force and indicate the bar's magnetic field.

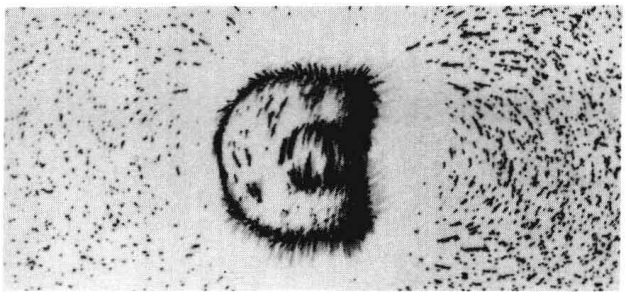


Figure 4.3 Horseshoe Magnet Iron filings indicate the magnetic field surrounding the horseshoe shape.

Although no one has ever seen a line of magnetic force, we may visualize it as a continuous loop or ellipse that flows out of the north pole of a magnet and then extends through space to the south pole, through the mass of the magnet, and out through its north pole into space once more. Millions of these lines of force exist in the magnetic field of even a comparatively small magnet. Lines of force are always *continuous* loops or ellipses. There is never any break in their continuity. Lines of force have two other important qualities. First, like rubber bands under tension, they always tend to shorten when possible. Second, they tend to move away from other lines of force. (Similar magnetic poles or electric charges repel each other.)

This latter property helps to explain the attractive force between unlike magnetic poles and the repelling force between like magnetic poles. When unlike magnetic poles are near each other, lines of force flow from the north pole of one magnet, through space, to the south pole of the other bar magnet, through space, and to the north pole of the first magnet. As these lines of force (which actually loop through both magnetic poles) attempt to shorten, like rubber bands under tension, the unlike poles tend to draw together.

With two like poles opposite each other, the lines of force surrounding one pole tend to travel parallel with those surrounding the other pole. Since

lines of force tend to move away from other lines of force, their effort to do so pushes the like magnetic poles apart.

At this point we might recall Coulomb's statement on magnetic poles: Like poles repel, unlike poles attract, and the degree of the repelling or attracting force is a function of the strength of the magnetic fields. Coulomb also observed that the degree of repulsion or attraction of magnetic poles varies inversely as the square of the distance between them.

Theories of Magnetism

For many years the accepted explanation of how a substance is magnetized or demagnetized was that set forth in Ewing's Theory of Magnetism (Figure 4.4). Ewing, a British scientist, had proposed that a piece of iron, nickel, or cobalt was made up of millions of tiny basic magnets, each with its own north and south poles. In turn, each of these tiny magnetic systems, which he referred to as **domains,** was made up of groupings of molecules of the three magnetic elements. When they were all lined up in one direction, with all the north poles facing one way and all the south poles facing in the opposite direction, the net effect was cumulative. The result was a great strengthening of the total magnetic field and the formation of a very powerful permanent magnet. Ewing's theory found support because it was a matter of common observation that heating a magnet, pounding on it with a hammer, or subjecting it to other magnetic fields—all actions that would tend to break up the north-south alignment of the magnetic domains—invariably weakened the magnet.

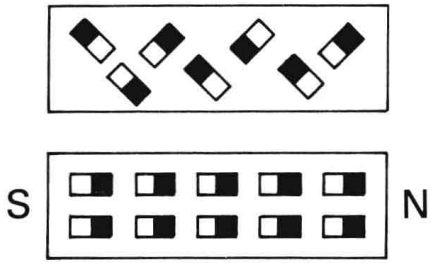


Figure 4.4 Ewing's Theory of Magnetism The iron bar on top is unmagnetized because the arrangement of its magnetic domains causes them to cancel each other out. The bar below is magnetized because its magnetic domains are all lined up in one direction and are reinforcing each other.

Present-day research explains the magnetic qualities of iron, nickel, and cobalt in terms of the behavior of the electrons that orbit around the nuclei of these elements' atoms. In nonmagnetic materials, researchers believe, electrons rotate in random directions as they orbit around the nucleus. However, in the case of iron, nickel, and cobalt, most of the orbital electrons rotate in the *same* direction. When an electron rotates, a moving magnetic field is also produced. Therefore, the effect of many orbital electrons rotating in the same direction becomes cumulative, and in effect the entire atom becomes a magnetic system with a north and a south pole. Scientists still support the concept of individual atoms linking together to form magnetic domains that when properly aligned, are the building blocks for the construction of a powerful permanent magnet.

Magnets as Complete Systems

It might appear that if we snapped a bar magnet in two, we would end up with two pieces of magnetic material, one with a north pole and the other with a south pole. However, as Figure 4.5 shows, we end up with *two* bar magnets, each with a north and a south pole. These two bar magnets can in turn be broken in two to produce two additional bar magnets, each with a pair of magnetic poles. Each time a large magnet is reduced into two or more smaller magnets, the strength of each individual magnet is reduced in direct proportion to its size.

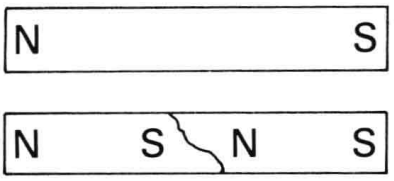


Figure 4.5 Breaking a Bar Magnet Every magnet is a complete system. When a magnet is broken in two (bottom), new poles are formed at the break.

Types of Permanent Magnets

Today permanent magnets of remarkable strength can be made from the various magnetic materials: iron, nickel, cobalt, and their alloys. One patented formula that has been in wide use for many years contains nickel, cobalt, and nonmagnetic aluminum. It is known under the trade name of Alnico. Alnico has gone through at least a dozen formulations since first offered to the electronics industry in the early 1930s. The addition of aluminum, which is basically a nonmagnetic element, appears to lock the domains of nickel and cobalt firmly in place once the magnetizing force has been withdrawn. (Apparently carbon serves the same purpose in magnets made of steel.) Alnico is still widely used in hundreds of electronic applications where a sturdy, very powerful magnet is required.

In recent years a new magnetic material has found wide application in electronics, particularly in situations where a very compact magnet with an intense magnetic field is desirable. The basic material is the metal cobalt, to which a precise amount of barium carbonate is added. Probably because of its physical appearance (hard and lustrous), magnets made of this material are referred to as **ceramic magnets.** For many applications Alnico magnets are still preferred, although they are appreciably more expensive.

Magnet Formation

Once the magnet structure has been properly formed mechanically, it must be energized. That is, the structure must be subjected to a magnetizing force that will line up the individual magnetic domains so that all the north poles face in one direction and the south poles in the opposite direction. The most effective way to energize a magnet is to subject the magnet's structure to a tremendously strong magnetic field for a very short length of time—perhaps a few seconds. The magnetic field is provided by passing a DC current of several hundred amperes or more through a coil of wire of comparatively few turns. The diameter of the wire must be at least that of a common lead pencil. After a second or two of current flow, the power is turned off and the magnet removed. It is now ready for use. While *permanent* is a relative term, tens of thousands of magnets manufactured in the 1930s and 1940s are still in daily service in electronic equipment without any apparent loss of efficiency.

Factors That Weaken Magnets

Permanent magnets such as we have discussed have found wide application in electronic devices like meters, loudspeakers, microphones, and high-quality phonograph pickups. However, certain factors can weaken or even destroy permanent magnets, resulting in meters that no longer give accurate readings or loudspeakers that fail to achieve normal volume. In general, anything that disturbs the arrangement of the magnetic domains will result in loss of magnetic efficiency—for example, applying heat; twisting, bending, or dropping the magnet; or pounding on it with a hammer. Less obvious, and therefore more dangerous, is the possibility of subjecting the magnet to stray magnetic fields—for example, placing a test meter on a workbench close to the power transformer of an operating TV receiver, or placing a dynamic microphone on a speaker cabinet containing a "woofer" speaker with a powerful Alnico magnet. Critical magnetic systems should always be kept at least several feet apart.

ELECTROMAGNETISM

Electromagnetic field: The magnetic field produced by the passage of a current through a coil.

The Danish scientist Hans Christian Oersted was the first to discover the presence of a magnetic field around a wire carrying an electric current. He noted that a nearby magnetic compass needle was deflected each time the current was turned on. The needle resumed pointing to the earth's north geographic pole each time the current was turned off. Oersted was puzzled at first by the fact that the compass needle did not point along the axis of the wire itself but rather pointed *toward* the axis of the wire. He theorized that the magnetic field around the wire was *circular*—that is, made up of an infinite number of lines of force arranged in concentric circles of varying diameters with the current-carrying wire as their axis. As the compass needle moved to permit these lines of force to flow through its mass, it actually was pointing in the direction of the lines of force rather than in the direction of current flow. In other words, the compass needle pointed at a right angle to the axis of the wire.

Later researchers found that forming the wire into a coil greatly strengthened the magnetic field for a given amount of current flow through the conductor (Figure 4.6). The behavior of an electromagnetic field is exactly the same as that of a magnetic field produced by a permanent magnet. The coiled, current-carrying wire is called an **electromagnet.**

Researchers decided that the increased strength of the electromagnet was caused by the lines of magnetic force surrounding any one turn linking with the lines of force of nearby turns. This linkage provided a greatly strengthened total magnetic field. Therefore, the symbol L (linkage) is used in electronics to represent a coil. Coils are referred to as **inductors.** Still later it was discovered that with the addition of a soft iron core to the coil, the same current flow and the same number of turns of wire would produce a magnetic field several hundred times more intense than when an air core was used.

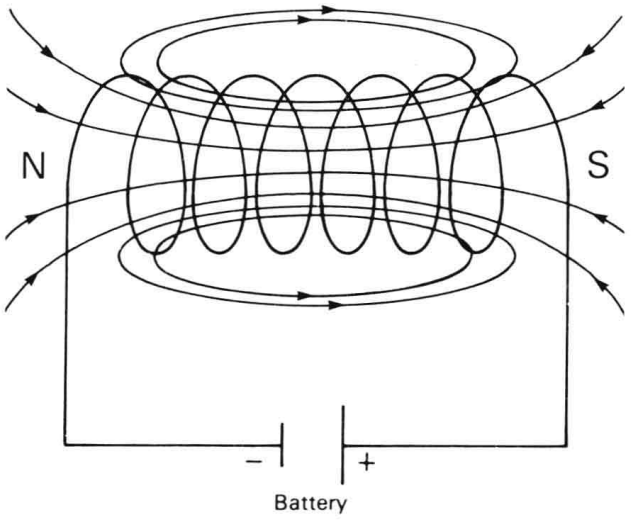


Figure 4.6 Magnetic Field around a Coil The coil shape strengthens the magnetic field for a given amount of current flow.

Early experimenters were greatly interested in producing an electromagnet with the greatest possible *lifting power*, or ability to attract objects made of iron or steel. In 1831 Joseph Henry developed an electromagnet with a lifting force of 750 pounds. Since insulated wire was not available commercially, Henry wound the electromagnet by hand, insulating the turns from each other by strips of silk torn from one of his wife's petticoats.

Strength of an Electromagnet

Researchers soon recognized that three principal factors determined the strength of an electromagnet: the diameter and the number of turns of wire in the winding, the nature of the core material, and the strength of the current through the coil. The same strength of magnetic field could be generated by two

different approaches: first, a very heavy current through a coil of a few turns of heavy wire and, second, a trickle of current through a coil of many turns of fine wire.

When direct current passes through a coil, the electromagnetic field around the coil behaves exactly the same as the magnetic field around a bar magnet. One end of the coil is always a north pole, the opposite end a south pole. Reversing the direction of current flow reverses the magnetic poles. If the coil is carefully suspended from its midpoint by a string, it will position itself to indicate the earth's north and south poles, just as a conventional compass needle does.

Applications of Electromagnetism

Today powerful electromagnets closely resembling the one Joseph Henry constructed in the 19th century are widely used to load crushed automobile bodies or steel scrap onto flatcars. A DC generator provides the heavy current needed to energize the electromagnet. When the load is to be dropped in position, the energizing current is either cut back or turned off.

The principles of electromagnetism are also applied in hundreds of other ways. For example, relays (magnetically operated switches) enable an operator at a radio or TV station to control potentials of thousands of volts with complete safety. Similarly, circuit breakers open an electrical circuit when too heavy a current flows. Motors and generators use two sets of electromagnets—one set turns (rotor) while the other remains stationary (stator). Other electromagnetic applications are far more complex. For example, recording audio and video information on iron oxide–coated acetate tape could not take place without electromagnetism. Storing and recovering information in the memory of a digital computer is another application of electromagnetics to modern electronics.

INDUCTANCE

Inductor: A coil that opposes any change in the value of current flowing through its windings.

Coils, or inductors, are very important in electronics, and their applications are remarkably varied. Some of the more common applications include the following: separating two AC currents of differing frequencies (e.g., separating radio from audio frequencies in a radio receiver); opposing any change in the value of current flowing through the windings (e.g., in the filter choke in a television set's power supply); and teaming up with a variable capacitor to produce a *tuned circuit* (the device that makes it possible for a radio or TV receiver to receive only the desired station and to reject all undesired stations).

How Inductance Works

One of the most important characteristics of a coil, or inductor, is that it always opposes any change in the value of current flowing through its windings. This property is called **inductance.** To see why this opposition takes place, refer to Figure 4.7. Here we have an inductor, L_1, in series with a battery. A switch is provided to turn the current through the coil on and off, and an **ammeter**—a device for reading the value of current flow—is included. When the switch is in the off position, the circuit is open, and no current can flow. When the switch is in the on position, the circuit is closed, and current can flow through the coil. We might expect the current through L_1 to reach its full value instantaneously, but it does not. Instead, the current starts out at rather a small value, rises rapidly, and finally reaches a steady value determined by the voltage of the battery and the resistance of the circuit.

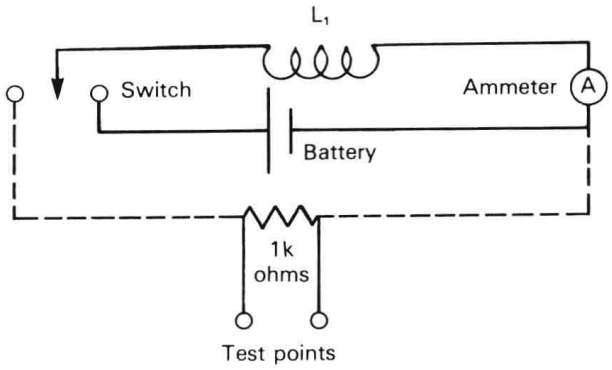


Figure 4.7 Inductor in Series with a Battery When the switch is turned on, current starts out at a small value and rises rapidly.

When the switch is moved to the off position, we might expect the value of current read by the ammeter to drop to zero immediately, but it does not. The current starts to fall immediately but does not reach zero until an appreciable length of time has passed. The fact that the current doesn't reach zero immediately indicates the existence of an alternate current path that will complete the circuit, such as a 1K resistor. The alternate path provides a good test point.

To explain why the value of current flow can never reach its maximum at the instant the switch is closed or can never fall to zero at the moment the same switch is opened, we need to review how a magnetic field can be used to induce current in a coil. Earlier we noted that when a coil of wire was lying within a magnetic field, if either the coil or the magnetic field moved, a voltage would be induced in the coil and would set up a flow of current. Also, if the magnetic field moved away from the coil, current would flow in one specific direction. When the field moved toward the coil, the current would reverse direction. It doesn't make any difference whether the coil moves in relation to the magnetic field or the magnetic field in relation to the coil. However, one or the other has to move, or no current will be induced in the coil.

In Figure 4.8, the magnetic field does the moving, not the coil. When the switch is closed and current starts to flow through the winding, a rapidly expanding magnetic field springs into existence and sweeps across the turns of the coil. The magnetic field induces in the coil a positive or negative voltage of such polarity that it opposes the current flowing from the battery. (Expansion of a magnetic field is equivalent to motion.) For at least a century this opposing voltage has been known as **counter EMF.** It is most pronounced in the case of a coil of many turns wound on an iron core.

As the current slowly reaches its maximum value, the magnetic field slows its expansion and finally reaches a steady state. Since the coil is fixed in position and the magnetic field is stationary, no voltage is now induced in the coil to oppose the original direction of current flow. The value of current that does flow is dictated by the battery voltage and the circuit's resistance.

When the switch is moved to the off position, the current supplied by the battery necessarily falls to zero. However, the ammeter indicates that current continues to flow. The explanation is as follows:

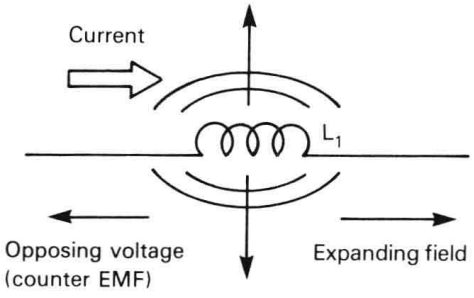


Figure 4.8 Production of Counter EMF The flow of current produces a rapidly expanding magnetic field around a coil, and the magnetic field induces an opposing voltage known as counter EMF.

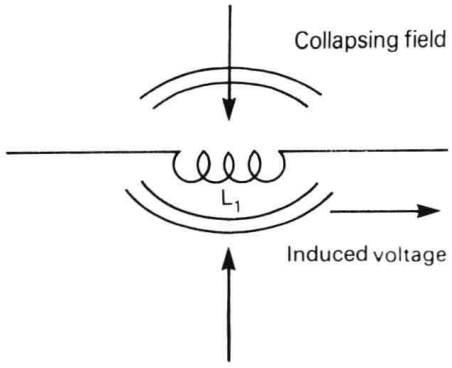


Figure 4.9 Collapsing Magnetic Field Turning off a switch causes the magnetic field to contract, which tends to keep current flowing until the energy stored in the magnetic field is returned to the circuit.

When the opening of the switch terminates current flow through the circuit caused by the battery, the fixed magnetic field around the coil rapidly contracts and, in so doing, induces a voltage in the coil (Figure 4.9). This time, because the field is contracting instead of expanding, the voltage induced in the coil is opposite in polarity to that produced when the switch was first turned on. Because the polarity of the voltage is opposite, it tends to keep the current flowing through the coil in the original direction after the battery voltage has been turned off. Current will continue to flow until all the energy stored in the magnetic field is returned to the circuit.

Inductive Reactance

Inductive reactance: The effect produced by a coil in opposing any change in the value of current flowing through its windings.

Inductive reactance comes about when the changing current through the coil produces a changing magnetic field around the coil. (When you react to something, you generally oppose it.) Inductive reactance is always expressed in ohms and usually is a great deal higher than the simple resistance of the wire contained in the coil's winding. A good illustration is given in Figure 4.10, which shows a simple series circuit made up of a miniature 110 volt lamp, used as an indicating device; a wire-wound resistor of 150 ohms; and an iron-core inductor, salvaged from a power supply, whose windings have a resistance of exactly 150 ohms, the same as that of the resistor.

When the test circuit is provided with a cord and a plug and is connected to an outlet providing 110 volt 60 hertz AC, the test lamp glows dimly. When an insulated test lead equipped with test probes shorts out the 150 ohm resistor, the lamp filament brightens appreciably but still remains at perhaps 20% of normal brilliance. If the test lead is transferred to the terminals of the inductor so as to short it out, the lamp brightens to about 80% of the brilliance it would normally provide when connected directly to 110 volt AC. Although both resistor and inductor have windings with the same measured resistance, the inductor is far more effective in restricting current flow than is the resistor. The explanation lies in the fact that with alternating current, the value of current flowing is constantly

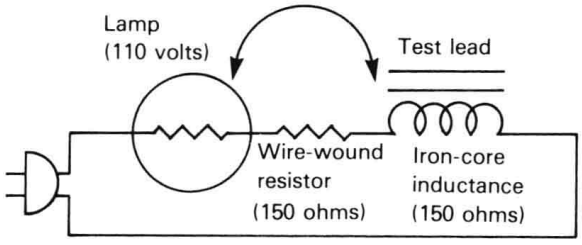


Figure 4.10 Inductive Reactance This simple series circuit demonstrates the effect of inductance on alternating current.

changing. Since the inductor opposes any change in the amount of current flowing, it is far more effective in opposing current flow than the simple resistance of its windings might lead us to conclude. Were this experiment to be repeated with 110 volt DC, shorting out either the resistance or the inductance would result in identical changes in the lamp's brilliance.

In electronics, the signals we most commonly encounter are either alternating current or pulsating direct currents. Signals that fall in the range of human hearing—say, from 50 to 15,000 cycles per second (referred to as 50 to 15,000 hertz)—are known as **audio frequencies.** For example, A above middle C on the piano is 440 hertz. Signals between 15,000 and 25,000 hertz are usually referred to as **ultrasonic waves.** Signals higher than about 25,000 hertz are known as **radio frequencies.** In general, a coil of a given inductance will offer decreasing reactance to low-frequency currents and increasing reactance to high-frequency currents. In Figure 4.11, two audio frequencies of equal intensity are seen entering a coil with an iron core. One has a frequency of 50 hertz; the other, 5000 hertz. The arrows on the right show their intensity as they leave the coil. The 50 hertz test signal has scarcely been changed in intensity. However, the 5000 hertz signal has been weakened so much by the coil's inductive reactance that it can barely be detected at the output. Two different frequencies can therefore be separated by the use of an inductor, a common application in electronics. If the

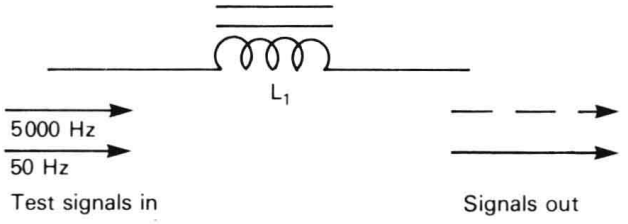


Figure 4.11 Effect of Frequency on Inductive Reactance The coil offers different reactance to 50 hertz signals than it does to 5000 hertz signals.

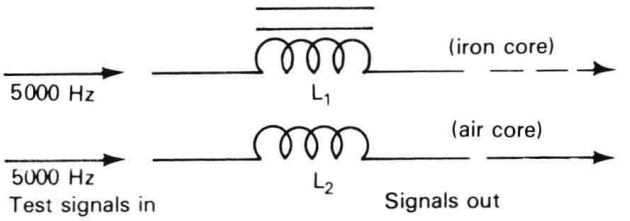


Figure 4.12 Inductive Reactance and Coil Type Air- and iron-core inductors offer different degrees of opposition to a signal of 5000 hertz.

test signal is of a constant frequency, it will encounter more inductive reactance in a coil with many turns and an iron core. It will encounter less inductive reactance in a coil with few turns and an air core (Figure 4.12).

Self-Inductance and Mutual Inductance

So far we have discussed only **self-inductance**—that is, the effects of the magnetic field surrounding a coil, or inductor, on the current flowing through the coil's windings. However, in electronic circuitry it is common practice to have two or more coils in close proximity. The changing magnetic field around the first coil reaches out and affects the current flowing in the second coil—a phenomenon known as **mutual inductance** (Figure 4.13). Mutual inductance is the basis of the transformer, discussed in a later chapter.

Solenoids

Solenoids are electromagnets with an energized cylindrical coil and an armature that moves back and forth within the coil. They are widely used in electronics and always involve a coil through which current is flowing. A simple sketch of a solenoid is shown in Figure 4.14A. When current flows through the coil L_1, the coil becomes an electromagnet with

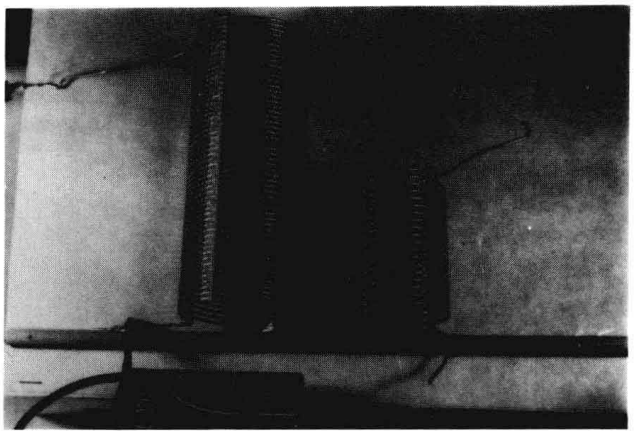
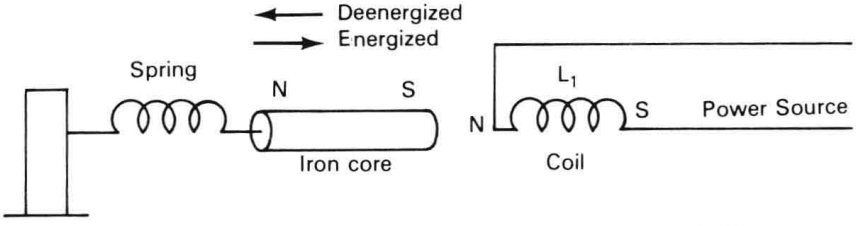


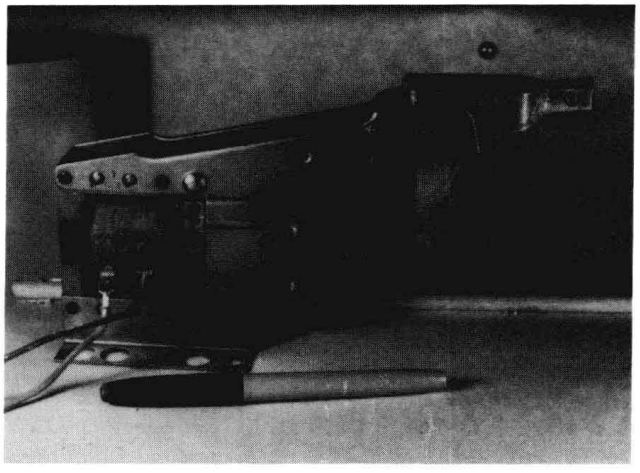
Figure 4.13 Mutual Inductance When a pulse of battery current is sent through the coil on the left, its expanding magnetic field induces a voltage and a current in the coil on the right.

one end a north pole, the opposite end a south pole. The field extending out from the north pole of the electromagnet automatically induces the opposite (or south) polarity in the near end of the iron core. The opposite end of the core becomes a north pole. The core should be made of soft iron since what the application requires is a temporary magnet.

With the north pole of the coil adjacent to the south pole of the iron core, a strong attractive force exists, and the iron core will move into the coil against the tension of the spring. This movement can be used to produce useful work, such as turning on or off a water valve in a washing machine (Figure 4.14B). When the current through the coil is turned off, its magnetic field disappears, and the tension of the spring withdraws the core to its original position. The solenoid is now ready to go through its cycle again.



A. The movement of the iron core toward the north pole of the coil can be used to produce useful work.



B. This solenoid, from a Westinghouse washing machine, shifts the drive mechanism from the wash to the spin-dry cycle. The left end of the spring attaches to an iron plunger; the right end, to an actuating lever. It is powered from a 110 volt AC source. (Pen shows scale.)

Figure 4.14 Schematic and Photo of a Solenoid These devices are widely used in the field of electronics.

Such a solenoid is also used in automobiles to turn on the starter motor, which draws up to 200 amperes at 12 volts. When you turn the ignition key, you close a switch, which energizes the solenoid's winding. (The solenoid is usually mounted on the case of the starter motor.) Energizing the solenoid closes a very heavy duty switch mounted nearby with contacts heavy enough to turn on or off a current of 200 amperes or higher without burning or arcing. Therefore, the ignition switch on the car's instrument panel needs to carry only a fraction of an ampere of current to control hundreds of amperes at a different location. Wide use is made of solenoids because they employ a low voltage and a small current to control high voltages or high currents from a remote location without hazard to the operator.

Measuring Inductance

The unit of inductance is the **henry,** named after Joseph Henry, a researcher in the fields of electromagnetism, telegraphy, and telephony. The henry is a satisfactory unit for measuring inductors with iron-core construction, such as filter chokes in power supplies and audio frequency transformers covering the range of 50 to 15,000 hertz. However, most coils found in radio receivers and TV sets have air-core inductors, and their inductance is more conveniently measured in millihenrys (one-thousandth of a henry) or microhenrys (one-millionth of a henry). By definition, a coil has an inductance of 1 henry when the change of current through its windings at the rate of 1 ampere per second will produce a voltage across its terminals of 1 volt. To simplify: A certain coil is carrying a current of 5 amperes. In one second's time the current is increased to 6 amperes. If this current change produces a voltage across its two terminals of 1 volt, the inductance of the coil must be 1 henry.

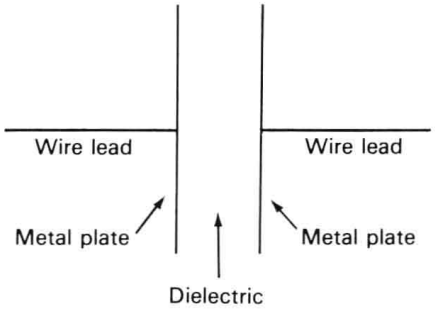


Figure 4.15 The Simplest Possible Capacitor The two metal plates are separated from each other by a dielectric (air).

CAPACITANCE

> **Capacitor:** A device capable of storing an electric charge in the form of an electrostatic field.

Next to resistors, the most common components on today's electronic circuit boards are **capacitors** (known until quite recently as *condensers*). In its simplest form, a capacitor, as shown in Figure 4.15, consists of two metal plates separated from each other by an insulating material known as the **dielectric.** Dry air is usually the standard dielectric to which all other materials are compared. Glass, mica, porcelain, and the various synthetics, such as polyethylene, are even more effective dielectrics.

Factors That Determine Capacitance

The factors that determine capacitance (the size of the charge a capacitor can store) are the area of the parallel plates, the spacing between the plates, and the type of dielectric material used. In general, increasing the area of the plates, decreasing the spacing between the plates, and changing to a more effective dielectric will all result in increased capacitance.

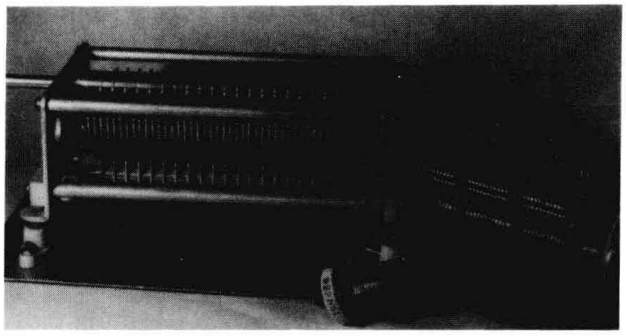


Figure 4.16 Three Types of Variable Capacitors A transmitting capacitor is on the left, a receiver tuning capacitor on the right, and a transmitter or receiver trimming capacitor in the middle foreground.

Variable Capacitors

While most capacitors are fixed in value at the time of manufacture, the capacitors used in tuning broadcast transmitters and radio receivers are frequently of the variable type and usually use air as the dielectric (Figure 4.16). As the tuning knob is turned, the *rotor plates* (attached to the turning shaft) mesh into the spaces between the *stator plates* (stationary). The rotor and the stator plates do not touch. They are separated from each other by the dielectric, air. A tuning capacitor has greatest capacity when rotor and stator plates are fully meshed and minimum capacity when the plates are completely out of mesh.

If an air dielectric capacitor, such as the tuning capacitor just described, is immersed in kerosene or mineral oil and then measured with a capacity meter, its capacity will have increased by a factor of 3 or 4 without any changes having been made in its construction. The increase is due to the fact that kerosene and mineral oil are more effective as dielectrics than is dry air.

Charging and Discharging Capacitors

In the case of an inductor, energy is stored in the form of an electromagnetic field surrounding the coil. The storage of energy by a capacitor takes place in an *electrostatic field* lying between the plates. Figure 4.17 shows the charging of a capacitor from a battery source. The strain exerted on the atoms of

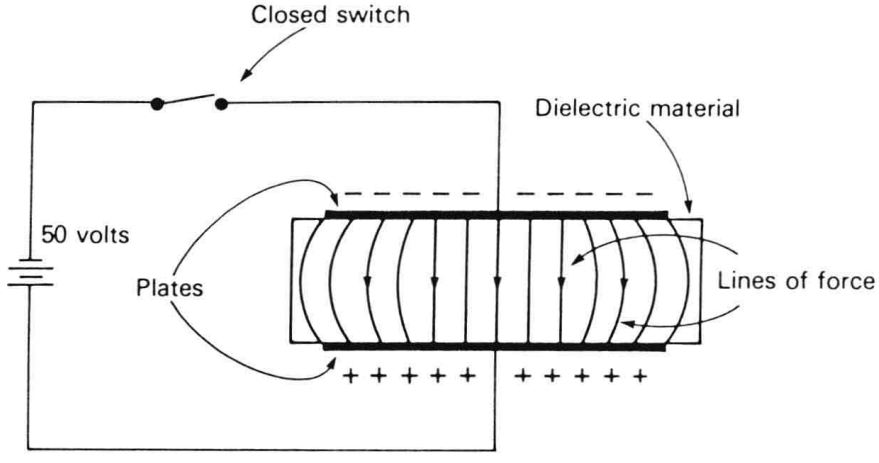


Figure 4.17 Electrostatic Field in the Dielectric Material of a Simple Capacitor Charging a capacitor creates an electrostatic field for the storage of energy.

the dielectric material by voltage applied across the plates produces electrostatic lines of force between them. Energy is stored in the resulting electrostatic field. It is believed that the positive nuclei of the atoms of dielectric material are displaced toward the plate with the negative charge. The negative electrons are displaced toward the positive plate. It is convenient to think of the stressed atoms of the dielectric in the same way we view rubber bands under tension—as a way of storing energy.

After the capacitor has once been charged and the switch connecting it to the battery is opened, we can use an insulated test lead to short one terminal of the capacitor to the other. For a fraction of a second we will see a hot spark, and then no further activity. We say that the capacitor has been *discharged*—that is, all the energy originally stored in the electrostatic field has now been returned to the circuit. In general, the greater the unit's capacity and the higher the voltage of the charging source, the greater will be the amount of energy stored and the hotter the spark will be on discharge.

Units of Capacitance

The unit of capacitance is the **farad** (named after Michael Faraday, who invented the electric motor, the generator, and the transformer). A capacitor has a value of 1 farad if, when charged to a potential of 1 volt, it yields 1 coulomb of electricity on discharge. A **coulomb** is the amount of electricity that passes a given point in a circuit during 1 second's time when the rate of the flow is 1 ampere. (The coulomb is not widely used in electronics.)

A capacitor of 1 farad would be so large that it would dwarf any circuit in which it was connected, so in practice much smaller values are used. Capacitors having values ranging from 0.01 microfarad to 1000 microfarads are frequently seen in power supply and audio circuits found in most receivers, amplifiers, and transmitters. A **microfarad** (mFd) is one-millionth of a farad. In the radio frequency circuits of the same pieces of equipment (frequencies above 25,000 hertz), values are much smaller and may range from 10 picofarads to perhaps 5000 picofarads. A **picofarad** (pF) is one-millionth of a microfarad.

Naming Capacitors

Capacitors are usually, but not always, named by the insulating material that serves as their dielectric. For example, there are air capacitors, such as the tuning capacitors in a radio receiver or transmitter; paper capacitors; capacitors similar in appearance to paper capacitors but with mylar or polyethylene as the dielectric; mica capacitors; ceramic capacitors using titanium dioxide as the dielectric; vacuum capacitors made of two metal plates sealed in a glass envelope and pumped to a high vacuum; and electrolytic capacitors. Figure 4.18 shows various types of fixed capacitors. Fixed capacitors, by far the most abundant kind, are those whose capacity does not change. Figure 4.19 shows various types of filter capacitors. Their function is to remove the "ripple," or voltage variation, from the output of a power supply. Filter capacitors generally have high capacity and high voltage ratings.

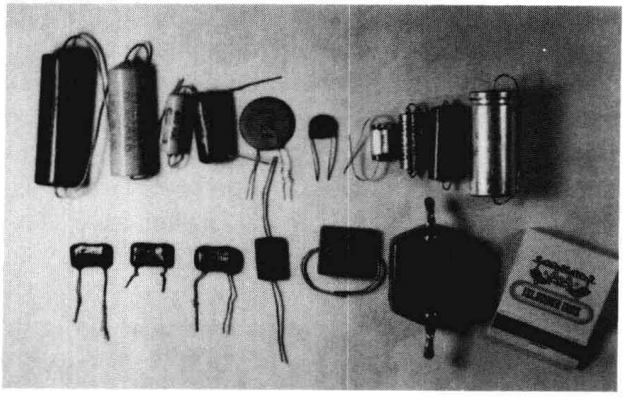


Figure 4.18 Various Fixed Capacitors Shown here are paper and plastic film (in upper left), mylar film (in upper center), electrolytic (in upper right), mica (in lower left), mylar film (in lower center), and disc-type ceramic (in lower right). The matchbook shows scale.

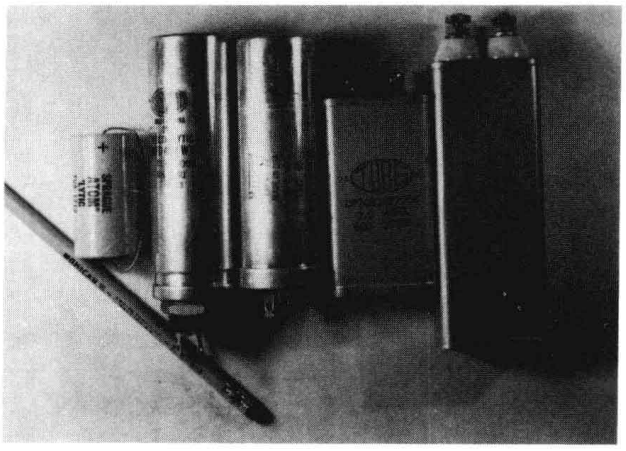


Figure 4.19 Various Filter Capacitors Far left: tubular electrolytic capacitors. Center: two electrolytic capacitors mounted in metal cans. Right: transmitting capacitors using oil-impregnated paper as the dielectric.

Electrolytic Capacitors

An **electrolyte** consists of an acid, base, or salt dissolved in pure water. A few pinches of ordinary table salt (sodium chloride) in a glass of water produces an electrolyte, or an electrolytic solution.

The electrolytic capacitor, of which there are many types, gets its name from the use of an electrolyte—usually sodium borate solution—to deposit a thin film of aluminum oxide on one of the capacitor's two aluminum plates during manufacture. Since the aluminum oxide is just molecules thick, the aluminum plates are effectively very close together. The result is a large amount of capacity in a very compact space.

To maintain the dielectric film of aluminum oxide on the proper plate at all times, the polarity of the capacitor's leads must always be observed. This polarity is prominently marked on the outside jacket of the capacitor. The lead designated + must connect to the positive side of the circuit, and the lead marked − must connect to the negative, or grounded, side. Failure to observe this precaution will cause the circuit to operate incorrectly and probably damage the capacitor permanently.

Since the chemical reaction that gave rise to the dielectric film of aluminum oxide at the time of manufacture must continue as long as the electrolytic capacitor is in service, such units always contain a certain amount of the original electrolyte. Consequently, their interior is always moist. If this moisture is lost, the capacitor will eventually lose all its capacitance. It is then said to be *open*.

Voltage Rating of Capacitors

Voltage rating: The maximum voltage that can normally be applied across a capacitor's terminals without causing an arc-over from one plate to the other or a rupture of the dielectric material.

Voltage ratings of capacitors vary widely, depending on the type of capacitor and the application for which it was designed. For example, electrolytic capacitors may vary from 5 to 450 volts in their voltage rating. Paper and mylar capacitors are widely used in the voltage range from 200 to 600 volts. Ceramic capacitors with voltage ratings as low as 100 and as high as 1000 volts are common, and mica capacitors intended for use in radio transmitters are readily available at working voltages of 5000 volts or higher.

Capacitors in Series and in Parallel

When manufacturers, electronics technicians, or experimenters require a capacitor of a higher voltage rating than the units available in the stockroom, they may decide to use two or more of the units on hand in series to provide the appropriate capacity and the needed working voltage. In Figure 4.20, C_1 and C_2 are each 2.0 microfarads capacity at a voltage rating of 1000 volts. When connected in series, the effec-

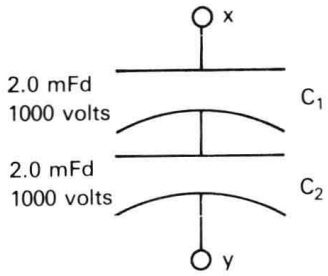


Figure 4.20 A Series Connection of Capacitors Connecting in series reduces the capacity but increases the working voltage. If C_1 and C_2 are each 2.0 microfarads at 1000 volts, the effective value (x to y) is 1.0 microfarad at 2000 volts.

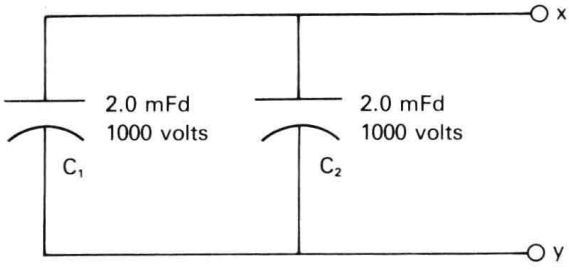


Figure 4.21 A Parallel Connection of Capacitors Connecting in parallel increases the capacity, but the voltage rating becomes that of the lowest-rated unit. The effective value shown here is 4.0 microfarads at 1000 volts.

tive capacity of the combination is halved, to become 1.0 microfarad, but the voltage rating of the combination is doubled to 2000 volts.

If you need 4.0 microfarads of capacitance at 1000 volts rating, you would connect two capacitors in parallel (Figure 4.21). The effective capacity becomes 4.0 microfarads at 1000 volts. Note that parallel connection increases the effective capacitance but has no effect on the voltage rating.

Precautions in Connecting Capacitors

When connecting capacitors in series, remember that they must be alike in both capacity and voltage rating. If unlike values of capacitance are connected together in series, the circuit voltage will divide across them unequally. The least voltage will appear across the terminals of the largest capacitor, while the most voltage will appear across the terminals of the unit with the least capacity. Under these circumstances, the smallest capacitor may "blow" and perhaps take important electronic equipment out of service. Also, electrolytic capacitors should never be placed in series with capacitors of other types, even if both have the same capacity. The electrolytic capacitor's characteristic leakage will produce an undue voltage strain on the other capacitor.

Capacitors may also be connected in parallel to provide additional capacitance at some specified working voltage. As seen in Figure 4.21, connecting two capacitors in parallel, each rated at 2.0 microfarads and 1000 volts, produces an effective capacity of 4.0 microfarads at a working voltage of 1000 volts. Any number of like or unlike capacitors can be connected in parallel without harm, but the voltage the combination can stand is that of the capacitor with the lowest voltage rating. The saying "A chain is as strong as its weakest link" applies to capacitors in parallel. For example, if a capacitor rated at 2.0 microfarads at 10,000 volts is placed in parallel with a capacitor rated at 2.0 microfarads at 100 volts, the effective value of the combination will be 4.0 microfarads at 100 volts. If a higher voltage than 100 volts is applied, the capacitor with the 100 volt rating may be destroyed and valuable electronic equipment damaged.

Capacitive Reactance

In our discussion of coils, or inductors, we said that an inductor always opposes any change in the value of current flowing through its windings. This opposition is known as inductive reactance. Capacitors

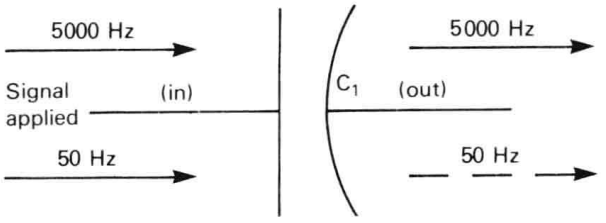


Figure 4.22 Capacitive Reactance The reactance of a given capacitor is lowest for high frequencies, greatest for low frequencies.

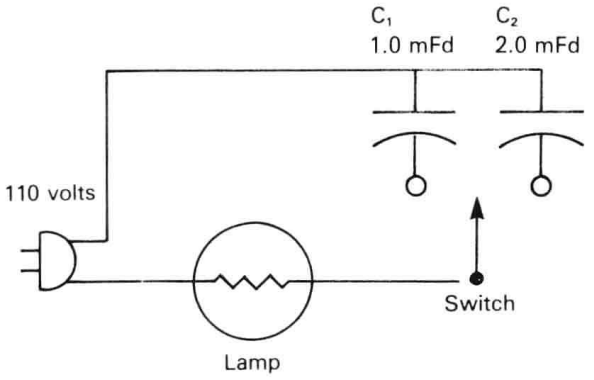


Figure 4.23 Reactance of Different Sizes of Capacitors The size of the capacitor affects AC current flow. The lamp is brighter when capacitor C_2 (2 microfarads) is in series with the AC line.

have a similar effect on the voltage in the circuit in which they are connected. This effect is called **capacitive reactance** and is always expressed in ohms.

> **Capacitive reactance:** A capacitor's opposition to any change in the value of the voltage in the circuit in which it is connected.

With inductive reactance we noted that for a coil of a given inductance, the opposition increased as the frequency of the current increased, and it decreased as the frequency of the current was lowered. Exactly the opposite is true of capacitors and capacitive reactance. As shown in Figure 4.22, a capacitor of a given size (C_1) offers little opposition to the passage of a 5000 hertz signal from one terminal to the other. At the same time it offers tremendous opposition to the passage of a 50 hertz signal, although both signals had the same original voltage intensity.

Figure 4.23 illustrates a situation in which the frequency of the voltage remains constant (60 hertz obtained from the wall outlet), but the size of the capacitor varies. As long as a capacitor is in series with the lamp and the 110 volt line, the lamp will not light to full brilliance. The capacitor represents an open circuit, which prevents a continuous flow of current. But as the capacitor charges first in one direction, then in the opposite direction as the AC voltage changes its polarity, enough electrons surge back and forth through the lamp filament to heat it and then light it. The larger the capacitance, the greater the number of surging electrons, as shown by the relative brilliance of the lamp filament. Using a smaller capacitor results in noticeable dimming of the lamp because fewer electrons are surging back and forth in the circuit. Fewer electrons produce less heating of the filament.

In an audio-frequency amplifier it was common practice to couple the signal in the plate circuit of one tube to the grid of a following tube with a capacitor. Today the collector of a transistor may be similarly coupled to the base of a following transistor. In both cases, a sufficiently large capacitor must be used so that the lower audio frequencies, particularly in the range 50 to 150 hertz, are transferred without appreciable loss.

COMBINING INDUCTORS AND CAPACITORS

In a number of applications, coils or inductors and capacitors are used together to accomplish some specific purpose. Figure 4.24 shows a typical power supply of about a dozen years ago that was used to provide high-voltage DC power for an audio frequency amplifier. The power transformer steps up the 110 volt potential from the AC outlet to about 300 volts. The rectifier tube then converts this AC voltage to pulses of direct current that flow in only one direction but vary widely in intensity. The two filter capacitors react with these pulses of DC to oppose the voltage variation, while the filter choke (an inductor) opposes current fluctuations caused by the voltage changes. Equipment manufacturers use these capacitors and the filter choke (an inductor) to minimize voltage and current variations. Without such devices, the pulses would cause an annoyingly loud hum in the loudspeaker that would drown out any voice or music.

Figure 4.24 Heavy-Duty Power Supply for Medium-Powered Audio Amplifier or Transmitter Upper left: power transformer. Upper center: duo-diode rectifier tube. Lower right: filter capacitors. Lower left: filter choke.

Tuning and Tuned Circuits

Another application of coils and capacitors is called a **parallel resonant circuit** (Figure 4.25). Parallel resonant circuits, widely used in receivers and transmitters, deserve brief mention here. (They are discussed in more detail in Chapter 7.)

With the inductance constant and the capacitor capable of being continuously varied in capacity by mere turning of the shaft, at any given setting of the rotor plates the circuit is "in tune" for some particular radio frequency. When it is in tune at a certain frequency, such as 1000 kilohertz (formerly 1000 kilocycles, or kc), the tuned circuit develops the maximum possible signal voltage at this particular frequency, and all other frequencies are weakened. Frequencies lower than the resonant frequency pass off to *ground* through the coil; those higher in frequency, through the capacitor. Only the resonant frequency is not weakened by the tuning process.

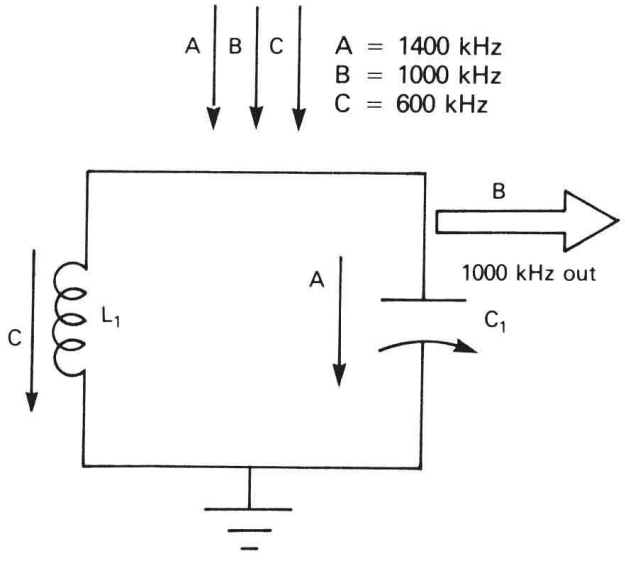


Figure 4.25 Operation of a Parallel Resonant Circuit L_1C_1 are resonant (in tune) at 1000 kilohertz. This frequency is strengthened, and all others are weakened.

Ground: The reference point in a circuit that is at zero signal potential.

Tuning makes possible the characteristic of a radio receiver called its **selectivity**—its ability to select just one station and reject all unwanted signals. A typical transistor receiver will employ about 6 parallel resonant circuits (tuned circuits) to separate stations in the standard broadcast band, which extends from 1600 kilohertz to 540 kilohertz.

Calculating Inductance and Capacitance

When inductors, or coils, are connected in series (without magnetic coupling between units), the result is an *increase in inductance*. The same formula is used as for resistors in series. When coils are connected in parallel, the result is a *decrease in inductance*. For coils in parallel, the same formula is used as for resistors in parallel (assuming no magnetic coupling).

When capacitors are connected in series, the effect is a *decrease in capacitance*. The same formula is used as for resistors in parallel. If capacitors are connected in parallel, the effect is an *increase in capacitance*. The formula used is the same as for resistors in series.

SUMMARY

Magnetism refers to the characteristics of iron, nickel, cobalt, and their alloys and the manner in which they are used to produce both temporary and permanent magnets. Every magnet has both a north and a south magnetic pole as well as an area of magnetic influence known as a magnetic field. The earth itself is a tremendously powerful magnetic system. Its magnetic field extends tens of thousands of miles out into space.

Electromagnetism describes the magnetic effects produced when an electric current passes through the turns of the windings of a coil, or inductor. The strength of the magnetic field produced is a function of the number of turns of wire in the inductor, the nature of the core material, and the strength of the current flowing through the winding. Magnetic fields of very great strength can be produced in this manner, and they behave exactly as do the fields surrounding permanent magnets. However, electromagnetic fields exist only as long as a current is flowing through the turns of the coil and disappear when the current no longer flows.

A coil opposes any change in the value of current flowing through its windings. This property is called inductance and is caused by the voltages and currents produced in the winding by the expanding and contracting magnetic field. The opposition the coil offers to the changing current through its windings is called inductive reactance and is expressed in ohms. When the magnetic effects around an inductor are confined to that coil alone, the reactance is referred to as self-inductance; when linked to a second coil, as mutual inductance.

By capacitance we mean that characteristic of a capacitor that permits it to store an electric charge in the form of an electrostatic field contained between two or more metallic plates. While inductors oppose any change in the value of current flowing, capacitors oppose any change in the value of voltage in the circuit in which they are connected. This opposition is called capacitive reactance and is expressed in ohms. In its minimum form, a capacitor consists of two metallic plates separated by an insulating material called a dielectric. Capacitors are usually named by the particular dielectric used in their construction.

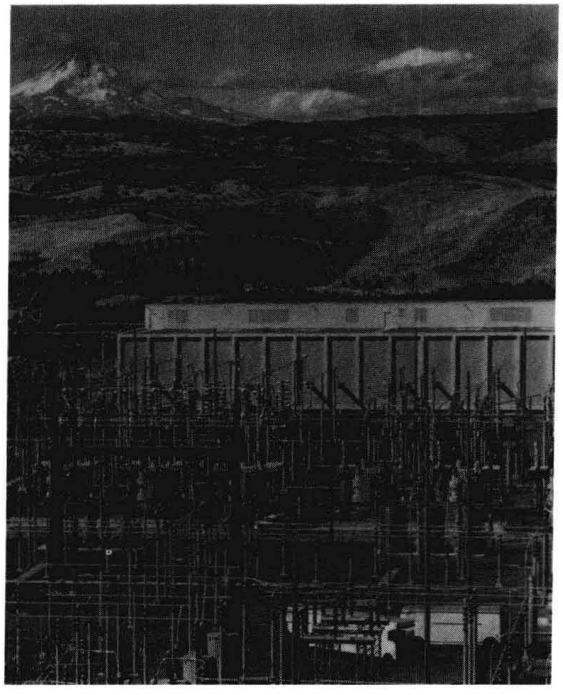
KEY WORDS AND CONCEPTS

alloy
ammeter
audio frequency
capacitive reactance
capacitor
ceramic magnet
coulomb
counter EMF
dielectric
domains
electrolyte
electromagnet
electromagnetic field
farad
filtering
ground
henry
inductor
inductance
inductive reactance
lines of force
magnetic field
microfarad
mutual inductance
permanent magnet
picofarad
radio frequency
retentivity
selectivity
self-inductance
solenoid
temporary magnet
tuning
ultrasonic waves
voltage rating

SELF-TEST QUESTIONS

1. Why is the earth's extensive magnetic field so important to its existing plant and animal life?
2. Which three elements have marked magnetic properties?
3. How can you distinguish between a permanent and a temporary magnet?
4. Describe a simple way of changing an unmagnetized bar of steel into a strong, permanent magnet.
5. What precautions should you take to prevent a permanent magnet from weakening?
6. In your own words, define an electromagnet.
7. What are the three principal factors that determine the strength of an electromagnet?
8. How does an inductor (coil) oppose any change in the value of current flowing through its windings?
9. In your own words, distinguish between self-inductance and mutual inductance.
10. What is a capacitor?
11. What are the three principal factors that determine the capacitance of a given capacitor?
12. Inductors (coils) oppose any change in the value of current in the circuit in which they are connected. What does a capacitor oppose?
13. An inductor and a variable capacitor are connected to form a parallel resonant circuit. How can the resonant frequency of the circuit be readily changed?

5 Basic AC Theory and Devices



At this switching yard and terminal building at Celilo, Oregon, AC power is stepped up and down and fed over transmission lines to electrical users all over the Pacific Northwest. (Photo courtesy of Bonneville Power Administration, USDOE)

Some Applications of Alternating Current

An Overview of AC
- How DC Differs from AC
- AC as a Power Source
- How AC Flow Is Produced
- Commercial Generators and Alternators
- AC Voltage and Current Waveforms
- Improving Alternator Efficiency
- Peak Values and Effective Values
- Impedance

Electric Motors

Transformers
- Step-Up and Step-Down Ratios
- Structure of Power and Audio Transformers
- Applications of Transformers

Summary

Key Words and Concepts

Self-Test Questions

We ordinarily think of alternating current as something that lights and warms our homes and runs our appliances. We're also aware that industry uses a large amount of alternating current both for the primary power needs of manufacturing and for industrial lighting. Substantially 90% of industry's electrical power requirement is met by the use of alternating current. Less than 10% comes from sources such as batteries and direct current generators.

SOME APPLICATIONS OF ALTERNATING CURRENT

Alternating current: An electrical current that reverses its direction of flow at periodic intervals.

Essentially all alternating current power is produced by rotating a coil (or coils) with an iron core at very high speed in the field of an extremely strong electromagnet. This rotation requires a very large amount of mechanical power, which is supplied either by the energy of falling water (in the case of hydroelectric plants) or by high-pressure steam turbines (in the case of plants fired by nuclear fission or fossil fuels).

Power production and utilization is just one application of alternating current. Most of the signals with which we are constantly dealing in electronics are either alternating currents, alternating voltages, or pulsating direct currents or voltages that closely resemble AC. For our purpose, we can define a *signal* as a change in voltage or current value or frequency that provides meaningful information.

A signal is produced when you talk into the mouthpiece of a telephone. The sound waves set up by your vocal cords strike the carbon microphone, and the role of the microphone and its transformer is to send an alternating current down the telephone line, varying at a voice frequency rate. Similarly, when you pluck a string on your electronic guitar, a weak alternating current moves down the cable to the amplifier, and its frequency is the same as the resonant frequency of the guitar string. In the case of an electronic organ, when you depress the A key above middle C on the keyboard, you generate a weak alternating current having a frequency of 440 cycles per second, or 440 hertz.

In radio broadcasting, in television, and in point-to-point communication, high-frequency alternating currents act as carriers for voice, music, and picture information. When we receive mail, the letter is always enclosed in an envelope that acts as a carrier. We discard the envelope after its purpose—the safe delivery of the letter—has been accomplished. The same basic procedure takes place in a radio or television station. Program information is combined with a high-frequency alternating current having the same frequency as that assigned the station by the FCC (see Chapter 7). By the combining of both signal and alternating current, the latter has its power, or frequency, varied at a signal frequency rate.

For standard broadcast stations, the alternating current generated by the transmitter and fed to the antenna will fall somewhere between 540 and 1600 kilohertz (1000 hertz); for VHF television stations, between 54 and 216 megahertz (1,000,000 hertz); for an FM station, between 88 and 108 megahertz. At these frequencies the electromagnetic field produced by the alternating currents in the station's antenna is forced out into space at the speed of all electromagnetic waves, including light waves—approximately 186,300 miles per second.

When these waves strike the receiving antenna, they set up a weak alternating current that travels through the input circuit of the receiver. There, by a process called *detection* (discussed in a later chapter), the carrier wave is separated from the program material and discarded. The program information is then further processed and fed to a loudspeaker—or, in the case of a TV set, to both a loudspeaker and the picture tube of the receiver.

AN OVERVIEW OF AC

In this section we examine how alternating current is generated, precisely what we mean by frequency, the waveforms of a typical alternating voltage or current, the factors that determine how much voltage and current we can develop, and the concept of *impedance*, which roughly corresponds to resistance in a direct current circuit. Hereafter we use the accepted abbreviations for alternating current (AC) and for direct current (DC). One final observation: Though the tendency is to stress AC current flow, we must recognize that such current flows through a circuit only because an AC voltage is driving it.

How DC Differs from AC

You will recall that current flow in a DC circuit is made up of a progressive movement of electrons from the negative pole of the circuit, where they are in excess, to the positive pole of the circuit, where a deficiency of electrons exists. The progress of individual electrons through the circuit is only at the rate of perhaps a few inches per minute, but the wave motion their movement triggers takes place at the speed of light—186,300 miles per second. As shown in Figure 5.1, electron flow and current flow are always in the same direction, even though they may increase or decrease in intensity, or stop and then start again.

With conventional AC current, electron flow and current flow reverse direction periodically (Figure 5.2). Actually, the electrons merely surge back and forth in the circuit because the positive and negative poles of the circuit are constantly changing in polarity. These polarity changes take place so rapidly that as soon as free electrons start to move toward the positive pole, the polarity of the circuit changes before they can get under way. As the electrons start to move in the opposite direction, the polarity once more changes, and they resume their original direction. The net effect is that electrons in an AC circuit surge back and forth rather than flow in one direction. Yet movement of these surging electrons can provide us with heat, light, motion, and many other effects. In a 60 cycle circuit, the current flows in one direction for 1/120 second, then flows in the reverse direction for 1/120 second (Figure 5.3). The two pulses of current flow going in opposite directions constitute one **cycle.** There are 60 cycles in a second.

Let's look at DC current through *graphical analysis*—that is, a visual device in which we describe values of voltage and current along a vertical axis and describe time along a horizontal axis. In this

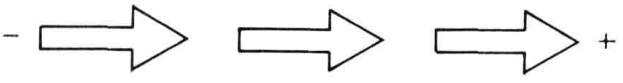


Figure 5.1 Direction of Current Flow in a DC Circuit Current always flows in the same direction.

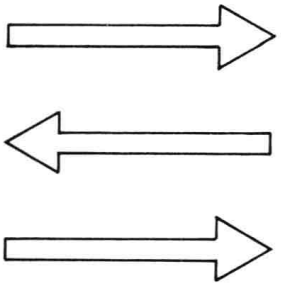


Figure 5.2 Direction of Current Flow in an AC Circuit Electrons surge back and forth as the polarity changes.

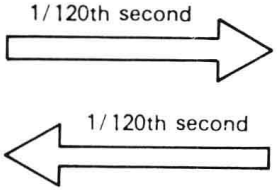


Figure 5.3 Current Duration and Flow in a 60 Cycle Circuit The two pulses of current flowing in opposite directions make up one cycle.

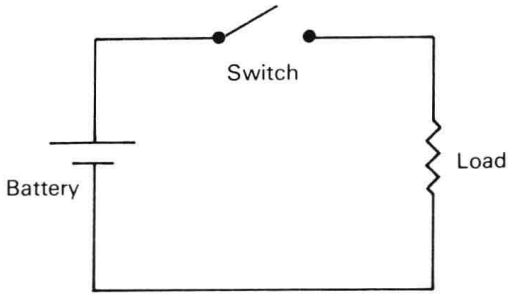


Figure 5.4 A Simple DC Circuit If the circuit is purely resistive, current will immediately rise to its full value when the switch is closed.

way, we can see more clearly what is actually taking place.

First, let's review what takes place in a DC circuit (Figure 5.4) at the moment the switch is closed. Assuming that the circuit is purely resistive and contains no inductance or capacitance, the current will rise to its full value immediately, as shown by the vertical line in Figure 5.5. Current will continue to flow at a constant level, as indicated by the horizontal line, until the switch is opened. At that time, current will again fall to zero. In essence, the graph has become a rectangle, and this is the way it will always appear when DC is turned on and off in a purely resistive circuit.

In interpreting such a graph, remember that time proceeds from left to right along the horizontal axis. "Most recent time" is at the extreme right of this line, while "least recent time" is at the extreme left. Along the vertical axis, voltage or current is zero at its intersection with the time axis and is increasing in amount above and below this point. It is conventional to assign the positive polarity to the vertical axis above its intersection with the time axis, and to assign negative polarity below this intersection. If the switch is opened and closed rapidly, like the movement of a telegraph key, a graphical plot of the results would look like Figure 5.6. Such movement is commonly referred to as **pulsating DC** (PDC).

Pulsating DC (PDC): A direct current whose steady flow is periodically interrupted.

PDC behaves very much like AC in electronic circuits, especially circuits that are handling signals. For example, when you talk into a crystal microphone, its output is a pulsating DC voltage varying at the same rate as your voice. (Its DC voltage varies in respect to ground as a reference.) Graphically, a pulsating DC voltage from an unvarying source appears as a series of rectangles. DC pulses of incredibly short duration and of varying voltage and spacing are the basis for the rapidly growing field of Digital Electronics.

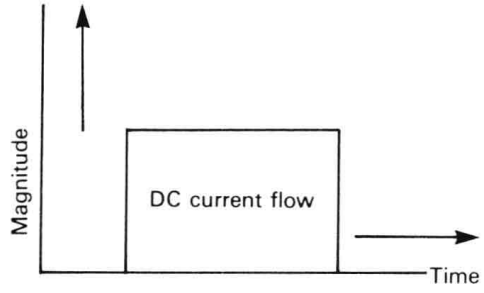


Figure 5.5 Graph of DC Current Flow Graphical analysis shows DC current flow as a rectangle in a purely resistive circuit.

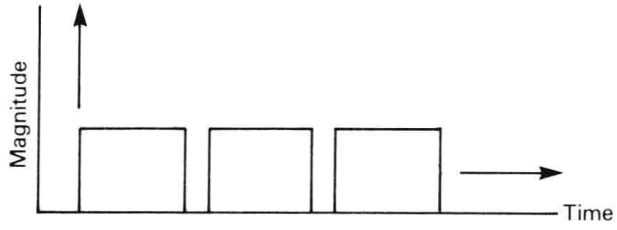
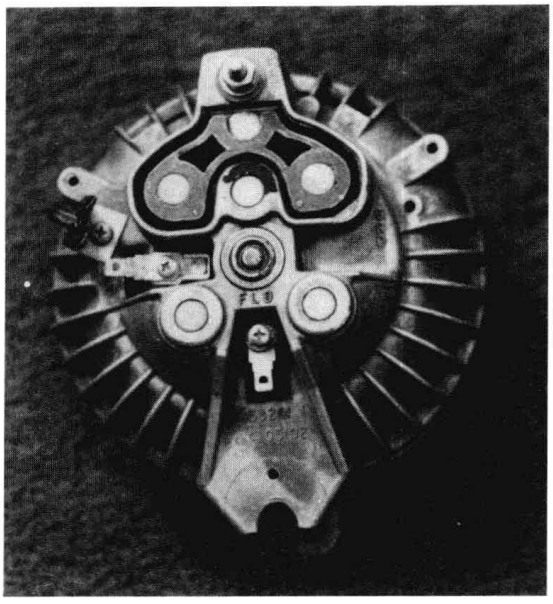


Figure 5.6 Graph of Pulsating DC Current Pulsating DC current results from repeated rapid opening and closing of the switch.



In this Chrysler Motors alternator from a Plymouth car, the three diodes used to change the AC to DC can be seen in a triangular pattern in the upper segment. The rotor is driven by the fan belt. (The diameter of the alternator is about 10 inches.)

AC as a Power Source

Most electrical and electronic equipment is powered by AC rather than DC. For every ten switches we throw in a typical business day, at least eight are probably connected in AC circuits. The other two would probably be the ignition switch on an automobile and perhaps the off-on switch on a pocket calculator or on a transistor radio.

Widespread use of AC as a power source can be traced back to Michael Faraday, who observed that when a coil of wire was moved toward a magnetic field, an electrical current would flow through the coil. When the same coil was withdrawn from the same magnetic field, a current would again flow, but in the opposite direction. Faraday then found it more practical to rotate the coil between a pair of magnetic poles. When he constructed this device in practical form, it worked very well indeed; Faraday had invented the alternator, or generator. He later invented the motor and the transformer.

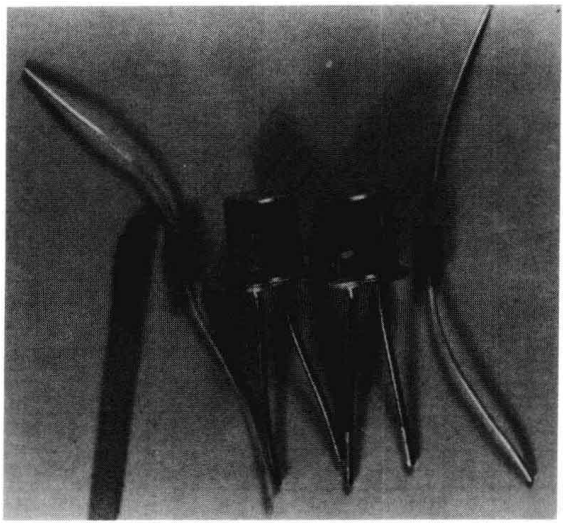
How AC Flow Is Produced

The electrons that make up the current flow produced by the generator or alternator come from the atoms of the copper wire in the coils of the generator's field and armature windings. The strong magnetic field pulls the single electron in the outermost shell of the copper atoms away as a free electron with a negative charge. The copper atoms that have lost electrons are no longer electrically neutral. They have a positive charge of 1 unit and are called *ions*.

In general, *alternator* and *generator* are interchangeable terms. Technically, however, an alternator delivers AC to its terminals through the medium of slip-rings and brushes. A generator delivers DC to its terminals through the medium of a commutator and brushes. All generators and alternators produce AC internally when the armature, or rotor, revolves in the strong magnetic field of the stationary field windings. The application for the specific machine determines whether this AC will be delivered as AC to the external circuit or first converted to DC.

Commercial Generators and Alternators

The AC generators at a hydroelectric or nuclear power plant are designed to work into the primary of a very large transformer whose secondary feeds the high-voltage transmission line that carries the energy to a nearby city or substation. There is no need to convert the alternator's output to DC, so it is directly connected to the step-up transformer (discussed in detail later in this chapter). The situation is different in the case of an automobile. Although the car's alternator is internally producing AC, DC is needed for several applications. For example, the storage battery must be charged with DC. The starting motor works far more effectively with DC than with AC,



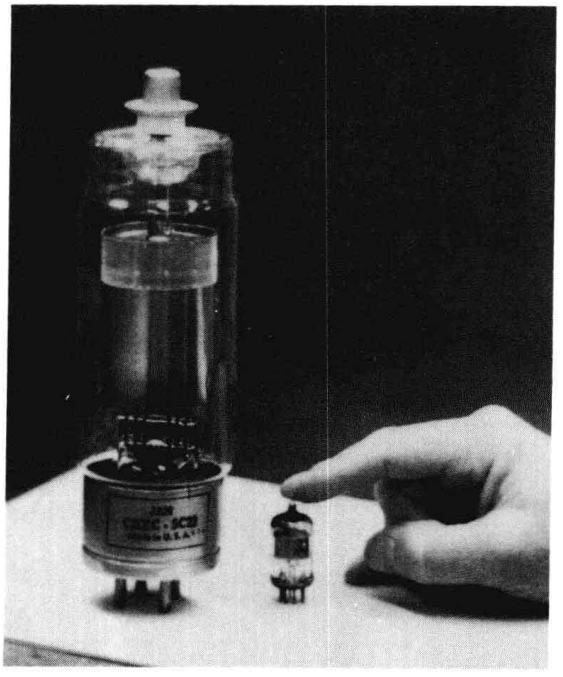
The scale of these typical solid-state diodes is shown by the paper match.

A diode can be considered a "one-way street" for electrons. Electrons can move from one terminal to the other terminal of the diode with ease when the current is flowing in one direction. When the current starts to flow in the reverse direction, the diode offers such extremely high opposition that no current flow takes place. Consequently, the output of an alternator equipped with one or more diodes is always made up of pulses of DC and always has current flowing in the same direction. Diodes have been in nearly universal use in the automotive industry since about 1955, and the power-generating device itself has been known as an alternator. Before 1955, all automotive generators used a commutator and brushes to convert internally generated AC to DC at the external terminals. Today solid-state diodes change the AC to DC, and the DC is carried to the external circuit by slip-rings and brushes.

and the ignition system would need to be redesigned completely if only AC power was available. So in this application, the alternator has its internally generated AC converted to DC before being brought out to its terminals. This conversion is done today by the use of solid-state diodes, which convert the AC generated within to pulses of DC in the external circuit. Usually three diodes are used.

Solid-state: An adjective used to describe any device using the crystalline form of a semiconductor like silicon. Free electrons are available by special chemical treatment of the silicon. No heating nor vacuum is required.

Diode: A device that conducts an electrical current readily in one direction and blocks its flow completely in the reverse direction.



The large vacuum-tube diode is used by the armed forces. The very small 6AL5 diode was widely used in older TV receivers.

AC Voltage and Current Waveforms

In Figure 5.7, we have what might be called an animated drawing. It shows exactly what occurs when a one-turn loop of wire (armature) is turned precisely one full revolution in a magnetic field lying between the north and south poles of a permanent magnet. Again, the horizontal axis represents time (the "most recent" time at the far right), and the vertical axis represents AC voltage. Increasing amounts of positive voltage are shown as you read *up* from the intersection of the two lines. Increasing amounts of negative voltage are shown as you read *down* from this point. The base line is calibrated from 0° to 360° because the loop of wire rotates exactly once. As the loop revolves, it describes one complete circle, and a circle is always 360°.

At the extreme left of the drawing, both horizontal sections of the loop are moving parallel to the lines of force that make up the magnetic field. No voltage is being produced in the coil, as the graph shows by having the trace remain at the zero point to the extreme left. (Remember to follow both the rotation of the coil and the tracing of the resultant curve from left to right.)

As the loop continues to turn, it starts cutting across the magnetic field at a steeper angle, and the voltage starts to rise. It becomes maximum when the loop is cutting across the magnetic field at a right angle (90°). Then the angle of cutting decreases, as does the amount of voltage developed in the rotating loop. When the loop is again moving parallel to the magnetic field, the voltage falls to zero (180°). As the loop continues to rotate, it starts cutting the lines of force at an increasing angle, and a voltage is induced within the loop. However, the two halves of the loop now have exactly the opposite relationship with the two magnetic poles that supply the magnetic field. Each half is moving closer to the pole that it was previously moving away from. Therefore, the polarity of the voltage induced in the

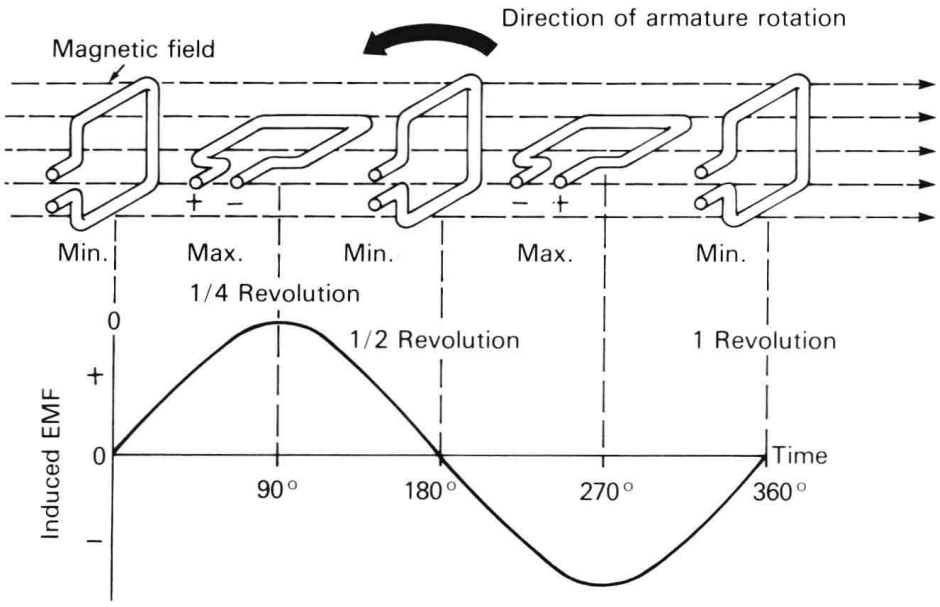


Figure 5.7 A Loop of Wire Turned One Full Revolution in a Magnetic Field AC current is generated and one cycle of a sine wave voltage induced.

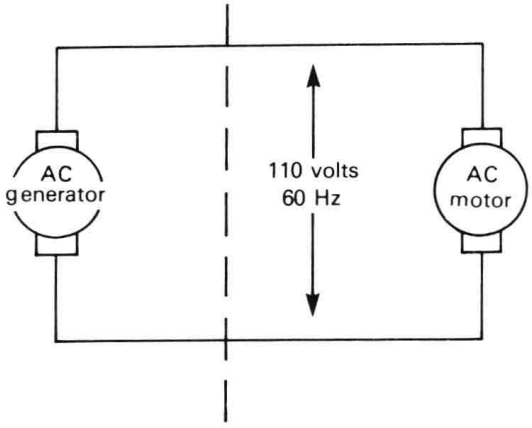


Figure 5.8 A Typical AC Circuit The generator is an AC alternator turned by water or steam. The motor is a typical unit powering a home appliance such as a food blender or a refrigerator.

coil is opposite to its original value and is now *negative*. This voltage rises to its maximum again at the moment it is cutting across the magnetic field at a right angle (the 270°). As the loop rotates its final quarter-turn, the angle of cutting becomes more and more shallow, and the voltage finally falls to zero at 360°.

The graph at the bottom of Figure 5.7 shows one complete AC cycle: From the original voltage of zero, voltage increases to a maximum in the positive direction, falls to zero, increases to a maximum in the negative direction, and falls again to zero. The waveform outlined in the graph is known as a **sine wave** and is sometimes referred to as the *signature of alternating current*. For each complete revolution of the loop, one cycle of a sine wave voltage is induced. (A typical AC circuit is shown in Figure 5.8.)

Pure and Distorted Sine Wave Tracings

When audio and radio frequency signals are completely free from distortion—that is, free from all **harmonics** (frequencies two and three times the fundamental frequency)—they exhibit the symmetrical waveform known as the sine wave. A common way of testing audio frequency amplifiers is to apply a sine wave having a frequency of 400 to 500 hertz to the input circuit and then view the resultant waveform in the output circuit with an **oscilloscope,** a test instrument very similar to a TV set. If the waveform seen on the oscilloscope is larger but still shows the same symmetrical form of the signal applied to the input circuit, the amplifier can be said to have low distortion. **Distortion** refers to any change in a signal (except an increase in its voltage or power) as it goes through an amplifier or other electronic device.

Standard Frequencies

In the United States, in Canada, and in Mexico, hydroelectric and steam-turbine plants that produce commercial AC power for distribution are standardized at 60 cycles per second (60 hertz), and this frequency is maintained with a high degree of accuracy. If the frequency were to increase to 61 hertz or decrease to 59 hertz, even for a short period of time, there would be serious consequences. Electric clocks would show incorrect time, and audio and video tape recorders would produce taped programs that could not be played over the air on broadcast and TV stations. Records played on the turntables of these same stations would sound too high or too low in musical pitch, and both computers and TV receivers might operate incorrectly.

While North America is pretty well standardized at 60 hertz for the frequency of AC power generation and distribution, in Great Britain and Western Europe 50 hertz AC power is universally used. Overseas, only nations with strong commercial ties to the West, such as Japan, use 60 cycle current.

Voltage and Current in Phase

Figure 5.9 shows the AC voltage waveform whose rotation for one revolution we described in Figure 5.7. If we connect this source of voltage to a resistive load, we obtain the waveform shown in Figure 5.10. Note that both AC voltage and current are shown

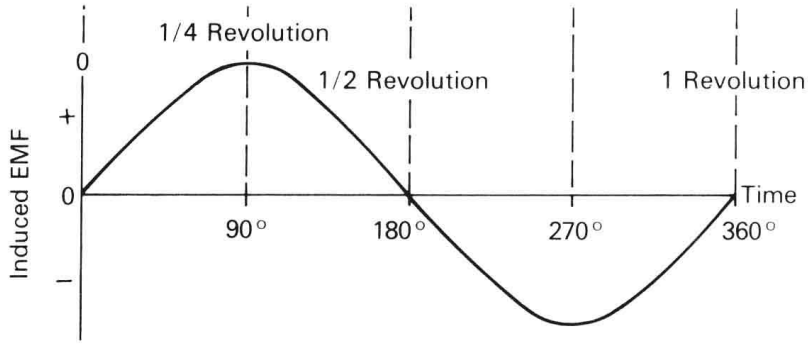


Figure 5.9 Waveform Produced by a Loop of Wire Turned One Full Revolution For each complete revolution of the loop, one cycle of a sine wave voltage is induced.

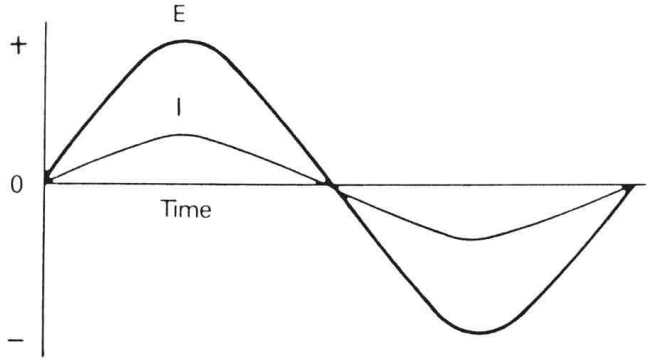


Figure 5.10 AC Voltage and Current in Phase Voltage and current reach their maximum and minimum values at the same instant.

because the current rises and falls in response to the rise and fall in voltage. As a result, voltage and current reach their maximum and minimum values at the same instant, and the E curve (voltage) and the I curve (current) essentially coincide. When both voltage and current are exactly in step with each other, they are said to be *in phase*. For all other conditions, they are said to be *out of phase*. For example, if there is a large amount of inductance in a particular circuit, the current changes will lag behind the voltage changes, which will have the effect of displacing the current curve to the right of the voltage curve. If capacitance predominates in a particular circuit, it will cause the voltage to lag behind the current and will displace the voltage curve to the right of the current curve.

Improving Alternator Efficiency

While the simple AC generator or alternator will actually generate a small AC voltage, the amount of power developed is insignificant. However, by comparatively simple changes in its design, its power output can be greatly increased. These changes are to: (a) provide a much stronger magnetic field by substituting for the permanent field magnet an electromagnet wound with many turns of wire over an iron core and energized by a DC current; (b) improve the effectiveness of the rotating coil (armature) both by increasing the number of turns in the winding and by winding it over an iron core; (c) provide additional armature coils mounted at varying angles to the original coil, providing pulses of power at the moment that the power supplied by the first coil is at its lowest value; and (d) provide for higher speed rotation of the armature coils in the intense field

supplied by the electromagnet. (In general, the more rapidly the armature turns, the greater its power output.)

Suitable means are required for leading the voltage and current generated within the alternator to the external circuits. Usually two insulated copper rings are mounted on the armature and connected to the armature coils. They are called *slip-rings* and are in constant contact with brushes of carbon or bronze that are suitably mounted on the alternator's frame. Wires lead from the terminals of these brushes to the external circuit.

Peak Values and Effective Values

Because we often use 110 volt, 60 hertz current both at home and on the job, we should know what is meant by *peak value* and *effective value*. We should also know what effect these values have on components used in the circuit.

In Figure 5.11, we have a graphical representation of one cycle of the typical 110 volt AC, 60 cycle power available at baseboard outlets in the average home. The graph reveals that when the effective value of voltage is 110 volts, the peak value is far higher (about 155 volts). Also, the amount of power available for use varies widely during the cycle. The peak value of voltage is comparatively high because the fluctuations of AC made it necessary to arrive at an effective value that would perform the same amount of useful work as 110 volts of DC in the same circuit. (In a DC circuit, both voltage and current may remain at a constant value for a long period of time.)

Recall that in an electrical circuit, the power in watts is equal to the voltage multiplied by the amperage. In Figure 5.11, as the cycle starts, both voltage and current are zero and the power is zero. At the peak of the positive cycle, the amount of power available is maximum. Then it decreases to zero as the curve crosses the horizontal axis. Next it rises to a maximum when the negative peak is reached and again slowly decreases to zero. Because the values of voltage, current, and power are varying over such wide extremes, the effective value of voltage is much less than its peak value. By international agreement, meters used for normal AC measurement are constructed and calibrated to read 70.7% of the peak value of the waveform. Again, this figure is

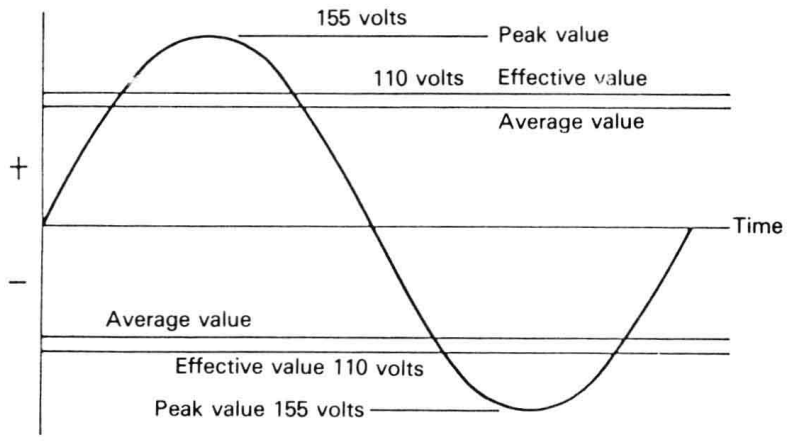


Figure 5.11 Analyzing the Characteristics of an AC Waveform When the effective value is 110 volts, the peak value is about 155 volts.

the *effective value* of the AC voltage and is sometimes referred to as the *root-mean-square value* (RMS value) because of the mathematical manner in which it was derived. If you know the effective value of AC voltage, you can always find the peak value by multiplying by 1.414.

Impedance

> **Impedance:** The total opposing effect offered by the resistance, the inductive resistance, and the capacitive resistance in a circuit.

By now you are familiar with the behavior of inductances (coils) and capacitances (capacitors) in electronic circuits where the values of both current and voltage are changing. By way of review, an inductance opposes any change in the value of current in the circuit in which it is connected. A capacitance opposes any change in the value of voltage in a circuit in which it is wired.

In an AC circuit the values of both current and voltage are constantly changing. In addition, therefore, to the opposition offered by the resistance in the circuit, there is additional opposition offered by inductances (called inductive reactance) and capacitors (called capacitive reactance). The word used to describe this total opposing effect is **impedance,** and its symbol is Z.

By substituting Z (impedance) for R (resistance) in various statements of Ohm's Law, its three basic equations can be used in AC circuit calculations. By way of review, these three statements are:

$$I = E/Z$$

$$E = IZ$$

$$Z = E/I$$

The actual formula for finding Z, or the circuit impedance, when the circuit's resistance, inductive reactance, and capacitive reactance are known is as follows:

$$Z = \sqrt{R^2 + (X_L - X_C)^2}$$

where R is resistance, X_L is inductive reactance, and X_C is capacitive reactance. (All three terms are expressed in ohms.) Since this text does not employ a mathematical approach to electronics, we will not work out a typical problem (Figure 5.12). However, note that when inductive reactance (X_L) and capacitive reactance (X_C) are both equal, they cancel each other out, and the circuit impedance becomes the same as the circuit resistance. Also, by way of review, recall that when capacitors and inductors are connected in series (or in parallel), they will always be in tune for some particular frequency. At the frequency at which they are resonant, or in tune,

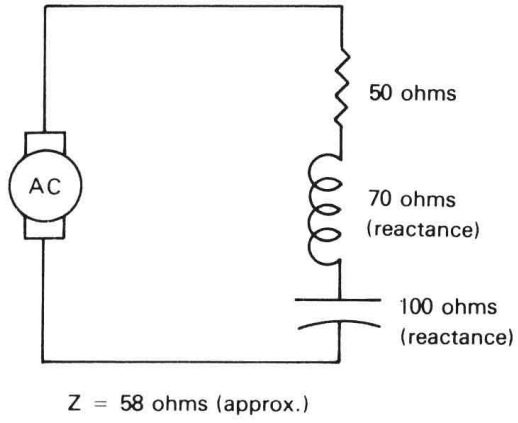


Figure 5.12 A Simple Impedance Problem AC circuits contain not only resistance but also inductive reactance and capacitive reactance, all of which oppose the flow of AC. The total opposition is called the circuit's impedance (Z). The calculated impedance (58 ohms) cannot be less than the resistance and often is much more.

capacitive reactance exactly equals inductive reactance, and the only opposition offered by the tuned circuit (coil and capacitor) at this one particular frequency is that of the circuit resistance alone. This opposition should normally be a very low value if the coil is wound with low-resistance wire and all connections to the tuning capacitor are securely soldered.

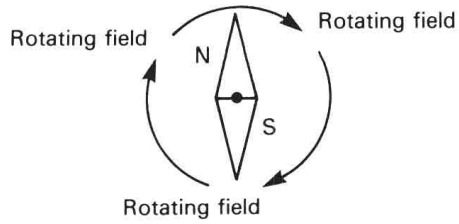


Figure 5.13 A Compass Needle Following a Rotating Magnetic Field The armature of the electric motor, like the compass needle, tries to follow the rotating magnetic field.

ELECTRIC MOTORS

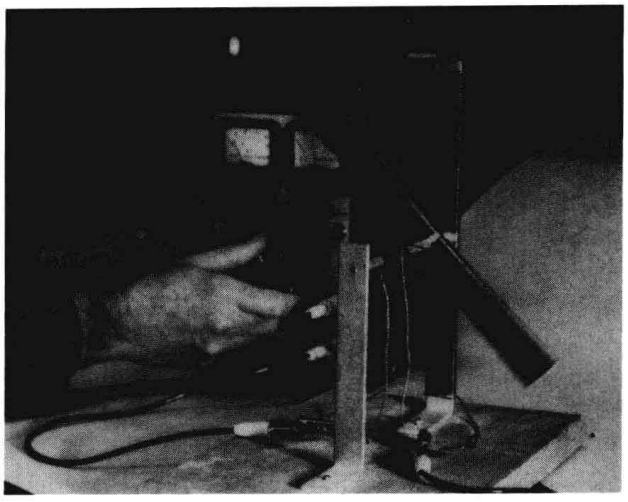
Since Faraday's invention of the electric motor in the early 1800s, the device has undergone tremendous development. The electric motor is closely related to the generator, or alternator, in that it has stationary field windings to provide a strong electromagnetic field and a revolving armature made up of many turns of wire wound over an iron core. The similarity between motors and generators is particularly true for motors equipped with both brushes and commutator. The commutator, made up of pairs of copper segments attached to the armature coils, is in constant contact with the brushes. It serves to commute, or reverse, the direction of current flow periodically. When the proper voltage is supplied to their terminals, both motors and generators will rotate and supply mechanical power. On the other hand, if they are supplied mechanical power to turn the shaft, and the field windings are energized from a DC source, a DC potential will be available from the terminals and can be used to operate other electrical and electronic equipment.

An interesting extension of this principle is seen by connecting the field winding in series with the armature. The motor will then operate either from an AC or a DC source. This type of motor is called a **universal motor** and is widely used in household appliances, such as electric mixers, blenders, and vacuum sweepers. Since some arcing always takes place as the brushes glide from one commutator segment to the next, they are always a source of some interference to radio and television reception.

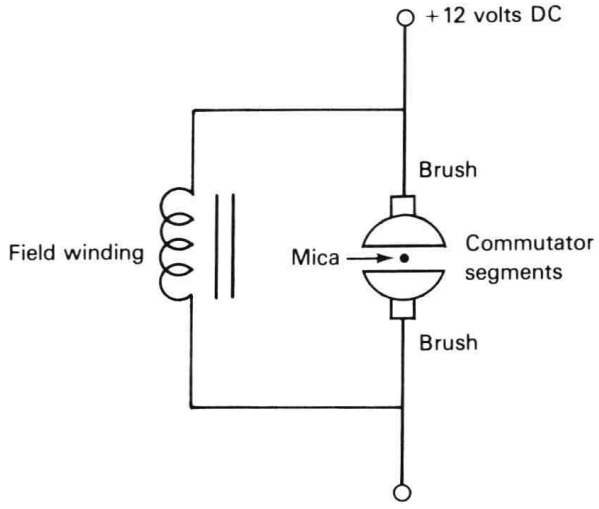
Basically a motor operates on the principle illustrated in Figure 5.13, which shows a magnetic compass needle rotating as it tries to catch up with a rotating magnetic field. Although this analogy is somewhat oversimplified, an armature winding in an electric motor turns for precisely the same reason.

Figure 5.14A shows a very simple electric motor constructed from a kit. The vertical fixed winding in Figure 5.14B is the field winding, which is energized from a 12 volt DC source. The other winding, mounted on the shaft, is attached to two half-moon segments of copper called the *commutator* (the equivalent of a double-pole, double-throw switch). On these segments rest springy pieces of copper wire that serve as brushes.

When the 12 volt DC supply is turned on, current flows through both the field and armature windings. One end of each winding becomes a north pole, the opposite end a south pole. The armature winding, being free to rotate, will now turn so that its south pole is opposite to the field winding's north pole, and its north pole opposite to the field winding's south pole. But just as it swings to this position, the turning of the shaft causes the copper wires to glide from one commutator segment to the other. This movement reverses the current through the armature and also reverses its magnetic polarity, so the armature turns another half-turn. At this time the brushes and the commutator reverse the current direction again, and the whole process is repeated. The net



A. The field winding is the vertical coil at the rear. The rotor, or armature, can be seen in motion. The commutator, in the foreground, reverses current direction through the armature to maintain turning motion.



B. The motor in this diagram is called a shunt field motor because the field windings are connected in shunt, or parallel, with the brushes.

Figure 5.14 Photo and Schematic of a Simple DC Motor

result is high-speed rotation of the armature winding upon its shaft. The poles of rotor and field never quite have time to line up opposite each other. If they did, the motor would stop.

TRANSFORMERS

Although Faraday invented the transformer early in the 19th century, it was first publicly demonstrated by George Westinghouse at the Chicago World's Fair in 1893. Westinghouse was displaying the world's first system for the generation and distribution of AC power, the same system almost universally used today. In doing so, he used transformers to step voltage up and down as the application required. A **transformer,** as the word implies, transforms electrical energy at one voltage level to a similar amount of energy at another voltage level. To accomplish the transformation, the device must have as a minimum two windings with magnetic coupling between them. The winding to which power is supplied is called the **primary.** The winding from which power is taken is known as the **secondary.** Frequently, a transformer has several secondary windings, but usually only one primary.

Step-Up and Step-Down Ratios

If the secondary winding has exactly the same number of turns as the primary, we say that the *turns ratio* is one-to-one (1:1). Therefore, 110 volts applied to the primary would result in 110 volts in the secondary. If the secondary had only half the number of turns as the primary, the secondary voltage would be 55 volts. If it had twice as many turns as the primary, the secondary voltage would be 220 volts.

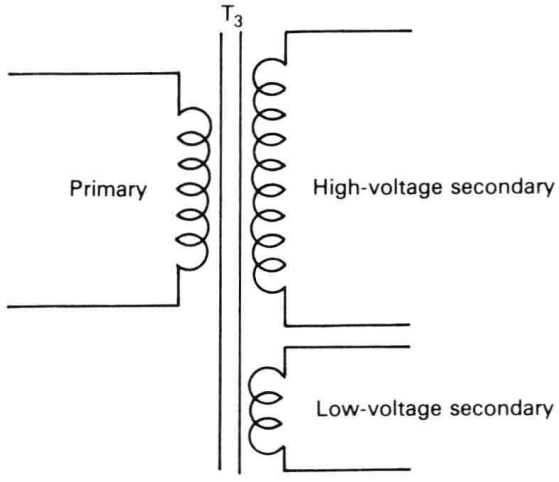
If you step up the secondary voltage by a factor of 5 (550 volts), you step down the amperage in the same winding by the same factor. In other words, you would have 5 times the voltage and 1/5 the

current. If the winding stepped down the voltage by a factor of 5, the secondary amperage would be stepped up by a factor of 5.

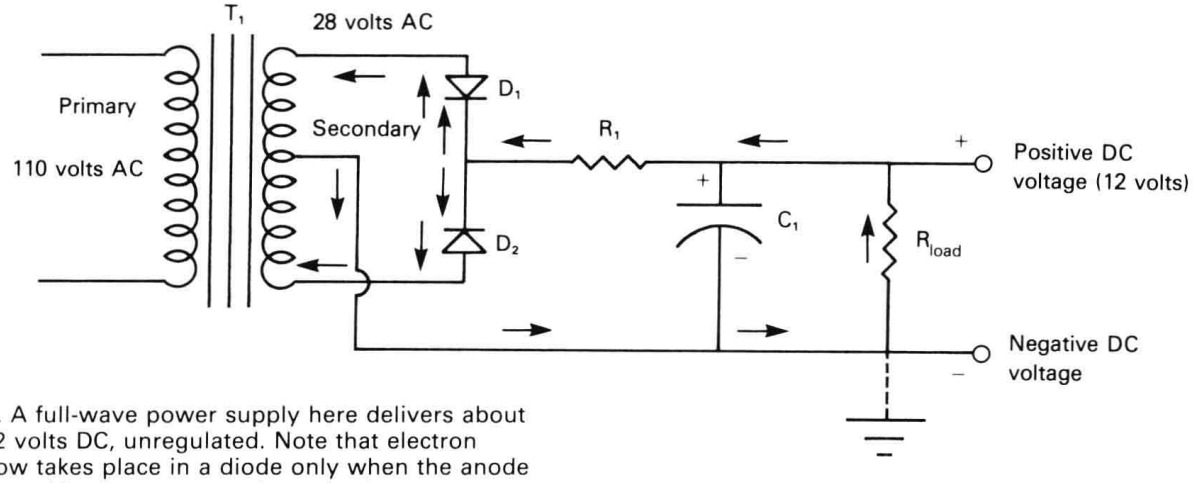
Structure of Power and Audio Transformers

The structure of power transformers and of audio frequency transformers is substantially identical. Although the core material is soft iron, it is made up of sheets of the metal, or **laminations.** Each lamination is insulated from nearby laminations either by varnish or by an oxide coating formed in the forging process. If a solid iron core were used, the changing magnetic field around the windings would induce a very heavy current throughout the core. The resultant overheating would damage the windings and waste a great deal of electrical energy. Figure 5.15A shows that a single transformer can be used to deliver simultaneously two different secondary voltages by using secondary windings of varying numbers of turns.

Let's assume that the transformer shown has 100 primary turns, a reasonable figure. The high-voltage secondary winding has 300 turns, and the low-voltage secondary winding has 10 turns. If 100 volts AC is applied to the primary, the high-voltage secondary winding will read 300 volts (3 × 100), and the low-voltage secondary winding will read 10 volts (1/10 of 100). There is no technical limit to how many secondary windings a power transformer may have, but the power in watts in all secondary windings added together cannot exceed the power in watts in the primary winding.



A. Step-up and step-down of voltage depends on the number of turns in the secondary windings.



B. A full-wave power supply here delivers about 12 volts DC, unregulated. Note that electron flow takes place in a diode only when the anode is positive in respect to the cathode.

Figure 5.15 Structure of Power and Audio Transformers

A Simple Power Supply

Many electronic devices, such as audio and video tape recorders, stereo systems, and low-power transmitters, are powered by a simple full-wave power supply as shown in Figure 5.15B. When 110 volt AC, 60 hertz is applied to the primary winding, approximately 28 volts appears across the center-tapped secondary winding. The ends of the secondary winding are connected to the anodes (positive electrodes) of two silicon diodes of a suitable voltage and current rating. Their cathodes (negative electrodes) are tied together and to the regulating resistor, R_1. R_1 is usually of the wire-wound type, and its resistance is determined by the voltage and current demands on the power supply. C_1 is an electrolytic capacitor of suitable voltage rating. Its capacity may be 1000 microfarads or larger. R_{load} is not an actual resistor but represents the current drain on the power supply caused by the electronic device connected to it (as well as by all the active devices and voltage dividers drawing this current).

R_{load} is determined by the following formula:

$$R_{load} = E/I$$

In the preceding formula, E is the output voltage of the power supply, and I is the current drain on the power supply. When AC is applied to the power transformer's primary, a lower AC voltage is applied to the secondary. As the cycle begins to change, the anode of D_1 is positive, while that of D_2 is negative. Half a cycle later the polarities change: D_1 is negative, while D_2 is positive. Strong electron flow takes place in each diode (cathode to anode) when the anode is positive, while no electron flow takes place when the anode is negative. (Arrows in Figure 5.15B indicate direction of electron flow.)

When viewed on an oscilloscope, the output of the diode rectifiers is a series of DC pulses, each with the same waveform of the 60 hertz AC wave originally applied to the transformer. The filter capacitor, C_1, opposes any change in the voltage applied to its terminals and converts this series of pulses (120 per second) to pure DC at the power supply's output terminal.

Electron Flow in a Power Supply

The flow of electrons through the power supply circuit gives rise to the positive and negative voltages at its output terminals. As each diode's anode in turn becomes positive, electrons flow from the diode's cathode to its anode, through half of the secondary winding of the power transformer, and from the center tap to the negative *bus,* or conductor. The electrons then flow along this line to the load resistance and through it to the positive terminal of the power supply. In so doing, the electrons produce a voltage drop equal to the voltage between the plus and minus terminals of the power supply.

From this point they flow along the positive bus, through the regulating resistor, and back to the cathodes of the rectifier diodes. The negative terminal of the supply may be (and frequently is) connected to the chassis ground of the electronic equipment in use. Note that the term *positive* is a measure of the *deficiency of electrons* at a certain portion in a circuit in reference to the normal, or grounded, state. In other words, +28 volts represents a certain level of electron deficiency measured in respect to the negative terminal, or ground.

Transformer Efficiency

The efficiency of a well-designed transformer intended for operation at power line frequencies can approach 95%. Transformers intended for use at audio frequencies (50 to 15,000 hertz) are somewhat less efficient since they must be designed to be equally effective over a wide frequency range. To obtain high efficiency, design engineers use as large a diameter copper wire as possible in the windings, wind the secondary windings directly over the primary windings to insure maximum coupling between the two, and use an iron core made up of laminations (thin sheets of iron) to minimize energy loss caused

by the changing magnetic field setting up circulating currents (**eddy currents**) in the iron core. Eddy currents serve no useful purpose. They waste electrical energy by converting it into heat.

At power line frequency (60 hertz) and at audio frequencies (50 to 15,000 hertz), transformers invariably have an iron core that produces the maximum electromagnetic efficiency with the fewest number of primary and secondary turns. Therefore, a very effective transformer can be of compact size. However, transformers designed for operation at radio frequencies (above 25,000 hertz)—for example, the tuning coils in a radio receiver—usually employ air-core inductors only. The use of an iron core increases the energy losses of the transformer at higher frequencies. In certain applications, however, a low-loss core of powdered iron may be used. Also, at radio frequencies the amount of inductance required is usually not so great as at power line or audio frequencies, so the iron core can be dispensed with.

When an AC voltage is applied to the primary of the transformer T_1, an AC current immediately surges back and forth within the coil and gives rise to an expanding and contracting magnetic field around it. This same magnetic field reaches out in its expansion and contraction and sweeps past the turns of the secondary winding, which is wound on the same iron core, to induce a voltage in it.

Applications of Transformers

Iron-core transformers have wide application in audio frequency circuits—for example, coupling an output stage in an audio amplifier or in a radio or TV set to a loudspeaker. Since transformers with good frequency response are expensive, they have been largely replaced in the best equipment by such circuitry as *complementary symmetry*, which permits a capacitor to couple audio frequency energy to a loudspeaker. However, much equipment still in service employs an output transformer for *impedance matching*.

To understand impedance matching, we need to look at the operation of a loudspeaker. A loudspeaker is basically an AC motor that operates at a very low voltage and a comparatively high current. It consists of a permanent magnet, a moving coil of just a few turns (a voice coil), and some means for attaching the terminals of the coil to a source of AC that is varying at a voice or music rate. When such a current is led into the coil, it sets up a varying magnetic field around it that reacts with the field of the permanent magnet. The reaction causes the coil to move back and forth over a short distance at a voice or music rate. Since one end of the voice coil is cemented to the cone of the loudspeaker, the fiber cone moves back and forth at a voice or music rate as well. In so moving, the cone sets up sound waves in the air, which the human ear perceives as voice or music.

The basic problem with this application is that the loudspeaker, as an AC motor designed to operate at a very low voltage and high current, must be energized by the 2N1711 transistor, which operates at a moderately high voltage and on a modest amount of current. In general, the resistance, or impedance, of the generator should approximate that of the load for effective power transfer. By using an output transformer (Figure 5.16) with a step-down turns ratio of about 9:2, excellent energy transfer from the output circuit of the transistor (about 80 ohms impedance) to the speaker's voice coil (about 4 ohms impedance) is obtained.

Another application of transformers is *isolation*. In radio and TV broadcasting, high-quality audio transformers are widely used to isolate one piece of equipment from another. For example, radio network programs coming into the station's control room over special, high-quality telephone lines are first fed to the primary winding of a transformer with excellent frequency response and approximately a 1:1 turns ratio. The transformer's secondary connects to the studio input equipment. This arrangement completely isolates the network service from the control room equipment and assures that a problem in one circuit will not have an adverse effect on the other.

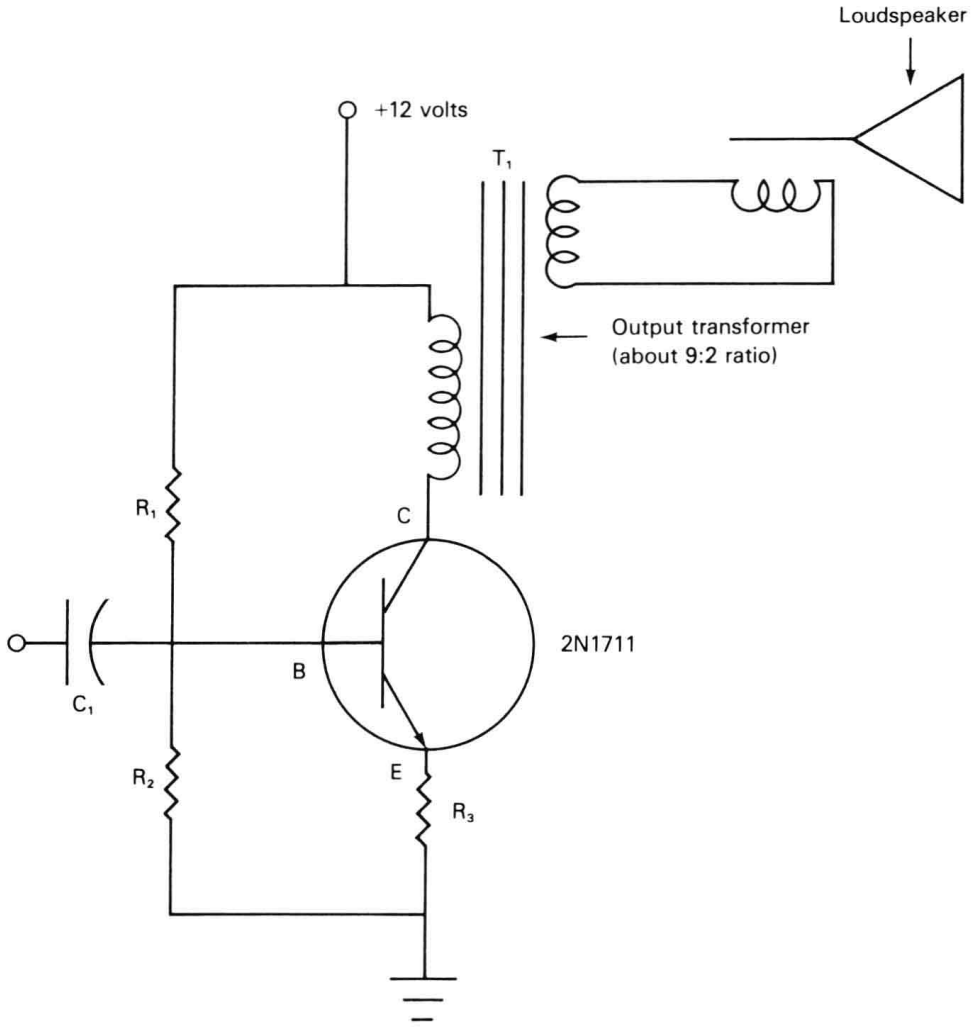


Figure 5.16 Output Stage of a TV Receiver The output transformer matches the approximately 80 ohm impedance of the transistor's collector circuit to the 4 ohm impedance of the voice coil of the loudspeaker. Impedance matching ensures effective transfer of power from the transistor to the loudspeaker.

Air-core transformers are widely employed in both receivers and transmitters for *tuning*. This process is described in detail in Chapter 7.

SUMMARY

Alternating current (AC) is considered the most important form of electrical energy for use in lighting, heating, and mechanical motion. The use of AC as a signal and as a carrier in electronic communications also deserves notice. The signal radiated by every transmitting antenna is actually an alternating current of a very high frequency of alternation. Even the output of a high-quality broadcast microphone or phonograph pickup head consists of a number of alternating currents varying at a voice or music frequency rate.

An alternating current is one that reverses its direction of flow at a periodic rate referred to as the frequency of the current. Normally frequency is expressed in hertz. Power is usually distributed by AC current having a frequency of 60 hertz in most of North America and 50 hertz in much of the rest of the world. Audio frequencies, a form of alternating current as they reach the loudspeaker terminals of your stereo set, fall in the range of 50 to 15,000 hertz. Signals radiated by standard AM broadcast stations range from 540 to 1600 kilohertz, and by FM stations, from 88 to 108 megahertz. VHF (very high frequency) TV stations include the frequencies between 54 and 216 megahertz.

There are two important reasons for the widespread use of AC for electrical power generation, distribution, and utilization. First, when hydroelectric or steam-turbine power is used to spin iron-core coils in an intense magnetic field, the process produces AC, not DC. (The production of DC would take one more step.) Second, and even more important, by the use of iron-core transformers, AC can be stepped both up (to a higher voltage and a lower current) and down (to a lower voltage and a higher current). Therefore, the voltage and current values best for a specific application can always be obtained. When a certain application requires DC voltages and currents (as in many vacuum tube and transistor applications), the AC is changed into pulsating DC by the process of rectification, which suppresses the negative half of the AC cycle and turns it into a positive pulse. A filter, made up of an iron-core inductance to suppress DC current variations and one or more capacitors that suppress DC voltage variations, converts the output into steady DC that compares favorably with the power provided by a battery.

The use of inductance and capacitance for removing voltage and current variations from the output of a power supply is based on their reactance to changes in current value and changes in voltage value in the circuits in which they are connected. Inductive reactance and capacitive reactance, always expressed in ohms of opposition, are very important in AC circuits since the values of current and voltage are constantly changing. For this reason, they must be considered along with resistance in determining the total opposition offered by a circuit to the flow of an AC current in that circuit. The proper term for the total opposition offered by the resistance, inductive reactance, and capacitive reactance to current flow in a particular circuit is impedance. By substituting impedance (Z) for R in the various statements of Ohm's Law, the equations can be used in solving for Z, R, and I when any two of these quantities are known in a specific AC circuit.

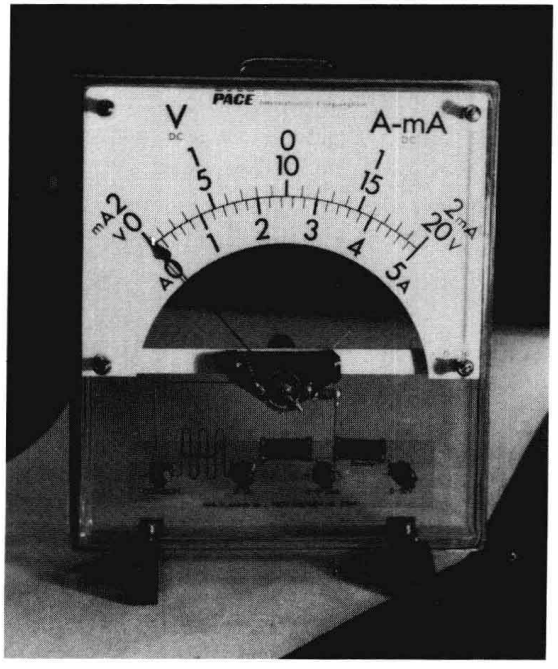
KEY WORDS AND CONCEPTS

alternating current
cycle
diode
distortion
eddy currents
harmonics
impedance
laminations
primary
pulsating DC
secondary
sine wave
solid-state
transformer
universal motor

SELF-TEST QUESTIONS

1. How is most AC power produced?
2. What other applications does AC have besides electrical power generation and distribution?
3. In your own words, describe how both voltage and polarity change during one cycle of AC.
4. How would you explain that while the effective value of conventional 60 hertz electrical power is only 110 volts, the peak value is 155 volts?
5. In addition to the opposition offered by resistance, the flow of AC current through a circuit is further impeded by two other factors. What are they?
6. How is the basic statement of Ohm's Law ($I = E/R$) modified so that it will apply to AC circuits?
7. Why does the armature (rotor) of an electric motor revolve when electrical power is supplied to its terminals?
8. Who is generally credited with the invention of the transformer?
9. Briefly explain why the same basic construction can be used for power transformers (handling 60 hertz) and audio transformers (handling 50 to 15,000 hertz).
10. A certain power transformer has 100 primary turns and 400 secondary turns. If 110 volts AC is applied to the primary winding, what will be the AC voltage across the secondary winding?

6 Basic Instrumentation and Storage Devices



This D'Arsonval meter is mounted in a Lucite case to reveal its components. In the center is the moving coil with attached pointer. On the lower left are two wire-wound meter shunts. On the lower right is the multiplier resistor for voltage range.

The D'Arsonval Meter
- Construction of the D'Arsonval Meter
- Operation of the D'Arsonval Meter
- Measuring Electrical Pressure
- Using the Meter on AC
- Measuring Current Strength
- Measuring Heavy Currents (AC)
- Connecting Voltmeters and Ammeters
- Measuring Resistance and Continuity

Volt-Ohm-Milliammeters
- Reading a VOM Dial Scale
- FET Voltmeters
- Digital Voltmeters
- Reading Power in a Circuit
- Measuring Power at Radio Frequencies
- Oscilloscopes

Cells and Batteries
- Carbon-Zinc Cells
- Rechargeable Cells
- Miniaturized Cells
- Charging Batteries

Summary

Key Words and Concepts

Self-Test Questions

For over a century and a half, scientists and experimenters have been asking such questions as: How much voltage is applied across this circuit? What is its resistance? How much current is flowing? Lacking suitable instruments to provide answers, they were forced to improvise. It was not uncommon to see an electrician of the 1930s inserting a calloused thumb in an electric socket to determine whether the circuit was live or using a 220 volt lamp with attached "pigtails" to trace out the wiring of an industrial control panel. If the bulb lit to full brilliance, it indicated that the test leads were across a 220 volt circuit. Half-brilliance showed that the wiring was connected across 110 volts. If insulated wiring felt warm, or even hot, to the touch, it was a good indication that current flow was excessive.

Early amateur stations frequently lacked even the most basic measuring equipment. In the 1930s, a station's sole piece of test equipment might have been a Readrite milliammeter purchased at the local Woolworth store. Full scale, it read 100 milliamperes (mils), so if the indicating hand slammed against the pin, it indicated that much more current was flowing. But there was no way of telling what the overload current was. Antenna ammeters were commercially available but not generally affordable. As a substitute, a Christmas tree lamp was connected between the antenna tuner and the antenna. The next step was to tune for maximum brilliance of the lamp. Though crude, this approach worked well. Stations were contacted and good signal reports received.

Today technicians have a wealth of test equipment (electrical measuring equipment)—for example, audio frequency oscillators, radio frequency signal generators and sweep-oscillators, color-bar generators, and distortion analyzers. Also available are the old standbys: the volt-ohm-milliammeter, the capacity meter, the tube and transistor checker, and the oscilloscope. Some of the latest volt-ohm-milliammeters, instead of having the familiar moving hand sweeping over a calibrated scale (analog meter), provide readings of voltage, current, and resistance in numeral form (digital meters).

In the following sections we examine how basic electrical and electronic meters were developed and what makes them work. The word *meter* comes from the Greek word for "measure." Meters are used for measuring voltage, current, resistance, and a number of other characteristics.

THE D'ARSONVAL METER

Although it will eventually be replaced, in part, by the new digital meters, one of the most widely used pieces of modern test equipment is the moving-coil meter known as the **D'Arsonval meter** (Figure 6.1). It consists of a high-quality permanent magnet of Alnico; a moving coil, or armature, positioned in the air gap between the magnetic poles of the permanent magnet; and a pointer, attached to a coil, that moves across a calibrated scale as the coil rotates. The moving coil is held in position by hardened steel pivots that rotate in watch-type jewels attached to the frame. Current is led into and out of the moving coil by means of helical (flat coiled) springs, which also serve to return the meter hand to zero when the measurement is completed.

The French scientist and engineer Arsene D'Arsonval, who developed this form of meter, is also known for his contributions to medical science. He was the first to suggest that victims of electrical shock could be saved if artificial respiration was started immediately. D'Arsonval was also one of the first to make a systematic study of the effect of radio and electrical waves on living tissue. His work laid the groundwork for what is known today as *diathermy,* the use of radio waves to produce heating deep within the body tissue.

Construction of the D'Arsonval Meter

D'Arsonval's meter, called a **galvanometer,** had a dial scale with zero current at its midpoint and increasing quantities of DC current on the scale both

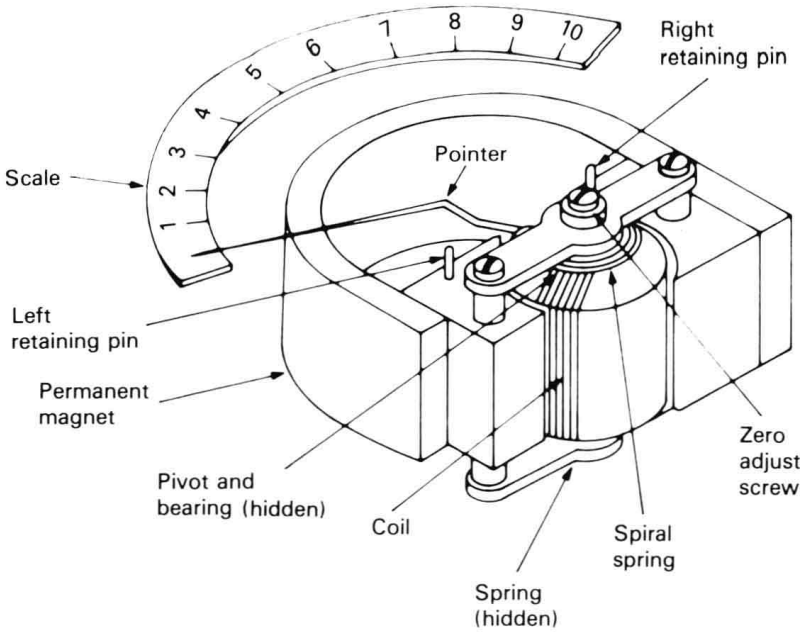


Figure 6.1 Movement of the D'Arsonval Meter A widely used piece of test equipment, the basic movement of the D'Arsonval meter is extremely simple.

to the left and to the right of the midpoint. It was not necessary to observe the polarity of the test leads since whatever the direction of current flow, a deflection either to the left or to the right of center-zero could be obtained.

D'Arsonval's original design has been modified by arranging the helical springs that carry current into and out of the moving coil so that they both tighten and relax at the same time. Therefore, zero position on the scale can be moved to the extreme left, which provides twice as many scale divisions between zero and full meter deflection and permits the meter to be read with much greater ease and accuracy. However, the polarity of the test leads connected to the meter must be observed at all times—usually by color coding the test leads (black for negative and red for positive). In this form the D'Arsonval meter is the basis for almost all meters used in electronic equipment today. The meter's basic construction has been the same for well over fifty years.

The use of hardened steel pivots and jewels to mount the moving coil between the pole pieces of the magnet is still the most common form of meter construction. However, in recent years manufacturers have used an alternate method of supporting the moving coil between the pole pieces of the magnet. The new method of support is known as the *taut band armature support* (Figure 6.2). Two short lengths of very light and flexible metal ribbon, usually with a 180° twist, extend from the meter's frame to the moving coil. When current passes through the coil, the taut band readily permits the coil to turn. When current flow ceases, the spring action of the ribbon of metal promptly returns the armature and the attached meter hand to zero. The taut band type of support has long been used in laboratory instruments where low friction and high accuracy were required. Only comparatively recently has there been substantial production of this type of meter for general use. The taut band type of meter is a true D'Arsonval meter, although of special mechanical construction.

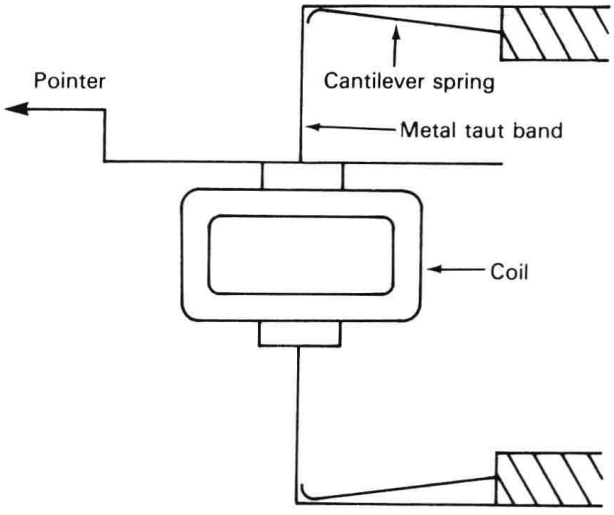


Figure 6.2 Taut Band Armature Support on a D'Arsonval Meter The taut band allows the coil to turn when current is flowing and promptly returns the attached pointer to zero when current is shut off.

Operation of the D'Arsonval Meter

Assume the pointer is resting over the zero mark on the dial scale, and the probes attached to the test leads are not in contact with any source of voltage or current. Since the meter sensitivity is not known, place a 1000 ohm resistor in series with one test lead to limit the current flow to a safe value. Now touch the test probes to a source of voltage, such as a flashlight battery (1 1/2 volts). The red test probe will contact the center terminal at one end of the battery; the black test probe, the bottom of the cell's case. Electrons will immediately start moving from the battery's negative terminal, up the negative test lead, through the moving coil, and back through the positive test lead to the other terminal of the battery.

As the electrons flow through the moving coil, an electromagnetic field springs into existence around the coil. This field reacts with the magnetic field supplied by the permanent magnet to produce a turning movement against the tension of the helical springs attached to the coil. At some point in the turning of the armature, the turning force becomes exactly the same at the opposing action of the tightening helical springs. As the coil stops its rotation, the pointer comes to rest over a scale division that the operator can read accurately. When one test lead is taken away from one of the battery terminals, current flow through the armature ceases, the electromagnetic field around it collapses, and the helical springs return both the moving coil and the attached pointer to the zero scale position.

Meter Damping

Early models of the D'Arsonval meter had one annoying fault: The indicating hand tended to oscillate back and forth like the pendulum of a grandfather's clock instead of coming promptly to rest over a scale marking. This defect was corrected in later models by winding the armature coil of copper wire over an aluminum bobbin, or support. When the armature starts to turn in response to current flow through its windings, the aluminum bobbin, moving in the magnetic field of the permanent magnet, has a voltage induced in it. The voltage gives rise to a current flow over its surface and to a weak electromagnetic field. This magnetic field around the aluminum form opposes any turning movement of the coil it supports. The net effect is the same as the action of a shock absorber on an automobile spring. The meter hand moves smoothly across the scale and stops quickly over the proper scale division without weaving back and forth. A meter of this type is said to be *magnetically damped*.

Factors That Affect Accuracy and Calibration

A number of factors affect calibration of the meter. The most important ones are the strength of the permanent magnet and the design of its pole pieces; the number of turns of wire in the moving coil and the design of the spool, or bobbin, over which it is wound; and the tension of the helical springs and the amount of friction in the armature's

pivots and jewels. In the case of taut band meters, design and tension of the bands themselves are also a factor. On a production line for meters, perhaps the most difficult single factor to control is the precise strength of the Alnico field magnet. The magnet is also perhaps the most vulnerable part of the assembled meter when it is in the hands of the owner. Subjecting the meter to a stray magnetic field, such as a nearby power transformer, can appreciably weaken the magnet and, in turn, noticeably affect the meter's calibration.

Meter Sensitivity

When a meter is advertised for sale, the literature invariably stresses that the unit is a "50 microampere movement" or perhaps a "1 mil movement." That is, a given meter will give exactly full-scale deflection when the advertised strength of current is flowing through its moving coil. *Exactly full deflection* means that just enough current is passing through the moving coil to bring the attached pointer to rest precisely over the maximum reading on the dial scale. A meter with 50 microampere movement will be far more expensive to manufacture than one with 1 mil movement since the moving coil on the former must be wound with hundreds of turns of wire smaller in diameter than a human hair—truly a job for a skilled technician. For most applications, however, the 50 microampere movement would be far more desirable since it draws just a trickle of current from the circuit under measurement and thus assures the highest possible accuracy. The less current drawn by the meter, the closer the measured value will be to the true circuit value—particularly in the measurement of voltage in circuits containing high values of resistance. In this application, the less current the meter draws, the more accurate the voltage reading will be.

Measuring Electrical Pressure

The basic D'Arsonval meter movement is limited to the measurement of very low values of voltage and current. In practice these low values are seldom encountered. However, by adding a few components such as resistors and batteries to the basic meter movement, the usefulness of the meter in making voltage, current, and resistance measurements of widely differing values is greatly expanded.

Figure 6.3 is a **schematic** (wiring diagram) that shows how the basic meter movement, with the addition of only four resistors, two selector switches, and a diode, can be transformed into a multirange voltmeter. This voltmeter will read voltages ranging from a minimum of about 1 volt to a maximum of 1000 volts with good accuracy. The addition of the solid-state diode makes it possible for us to read AC values with the same DC meter.

Assume that the meter in Figure 6.3 has a sensitivity of 1 milliampere. In other words, the indicating pointer will rest exactly over the 1.0 milliampere mark on the dial scale when 1 mil of current is passing through its windings. Although much more sensitive meters are used in commercial voltmeters, the meter shown will work well and is entirely practicable. Note that in one lead from the meter, a

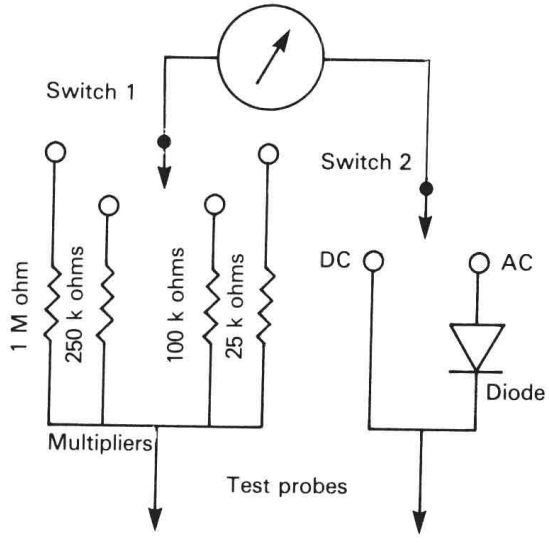


Figure 6.3 A 1 Mil Meter Connected as a Voltmeter The addition of resistors, switches, and a diode transforms a basic meter into a multirange voltmeter.

switch selects one of four resistors and, in effect, places them in series with one of the meter's test leads. Since the function of these resistors is to multiply the original range of the meter, they are known as **multipliers,** or **multiplier resistors.**

The resistance of a multiplier resistor is always greater than that of the moving coil of the meter. However, even the moving coil's resistance by itself can be used for very low values of voltage. A typical resistance for the moving coil of a 1 mil meter movement would be about 50 ohms. By substituting this value of current and this value of resistance in one of the statements of Ohm's Law ($E = IR$), we can solve for E and find the largest value of voltage that can be measured without an external multiplier. $E = 0.001 \times 50$. (In the preceding equation, 0.001 is 1 mil expressed in amperes. There are 1000 mils in an ampere.) The answer, 0.05 volt, is a value almost never encountered in any type of electronic equipment. Therefore, let's focus on more practical voltage ranges: 25, 100, 250, and 1000 volts, full-scale deflection.

The formula for finding the value of the multiplier resistor (R_{mult}) is very easy to use (internal resistance of meter disregarded):

$$R_{mult} = E_{max}/I_{meter}$$

In this equation, E_{max} is the highest voltage to be read on that particular range, while I_{meter} is the amount of current (in amperes) required to give exactly full-scale meter deflection.

EXAMPLE

Compute the value of multiplier resistor necessary to provide full-scale readings of 25, 100, 250, and 1000 volts with a 1 mil (0.001 ampere) meter movement.

Solution

Substitute and solve:

$R_{mult} = 25/0.001 = 25{,}000$ ohms (25 kilohms)

$R_{mult} = 100/0.001 = 100{,}000$ ohms (100 kilohms)

$R_{mult} = 250/0.001 = 250{,}000$ ohms (250 kilohms)

$R_{mult} = 1000/0.001 = 1{,}000{,}000$ ohms (1 megohm)

If you wish to use a meter at hand in this application, you can calibrate it against a commercial meter of good accuracy and mark the various values of voltage on the meter scale with india ink.

Using the Meter on AC

If the test leads of our meter are connected to a 110 volt, 60 cycle wall outlet (range switch set to 250 volt position), no meter indication will be obtained. The reason is that when the current from the outlet starts to flow through the meter's moving coil, it takes a certain amount of time for the meter to start moving. But the duration of the current pulse in one direction is only 1/120 second, so the armature only *starts* to move. Then comes a pulse of current in the reverse direction for 1/120 second. Again, the meter's moving coil just starts to move in the reverse direction in response to the reversal of current flow. The net effect of the rapidly changing direction of current flow is that the meter hand doesn't move noticeably, though it may vibrate slightly.

The solution is to change the AC flowing in the test leads to pulsating DC so the turning force exerted on the moving coil will be constantly in the same direction, rather than continuously changing. The conversion to pulsating DC is made through the use of a solid-state diode, an electronic device that looks like a small carbon resistor and acts as a one-way street for electrons. Electrons can flow only from the diode's cathode (negative terminal) to its anode (positive terminal)—never in the reverse direction. When AC is connected to the diode in our meter circuit, current can flow only during the positive half of the AC cycle. The current flow is blocked com-

pletely during the negative half. As a result, the diode's output consists of pulses of direct current varying in intensity but always flowing in the same direction.

When this pulsating DC is fed to the armature coil of the meter, a constant turning effort is produced, and the indicating hand comes to rest over some calibrated voltage point on the scale. Our meter will read a little low on the AC ranges since with a single diode, there will be only one pulse of DC lasting 1/120 second during one cycle, which lasts 1/60 second. In commercial meters, this problem is resolved by using four diodes in a bridge circuit to deliver a larger average amount of DC power to the meter movement. For further accuracy, a number of manufacturers also use separate multiplier resistors for the AC voltage ranges.

Measuring Current Strength

As is shown in Figure 6.4, the 1 mil D'Arsonval meter can be used to measure currents of various values. The meter needs no additional circuitry to measure currents ranging from near zero to a maximum of 1 milliampere. But when additional current is to be measured, most of it must be *shunted* around the meter because the meter coil must never carry more than 1 mil of current.

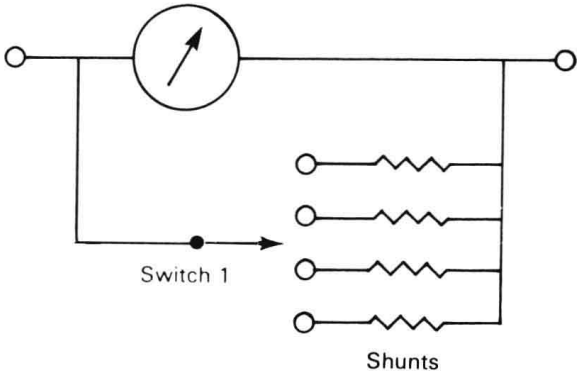


Figure 6.4 A 1 Mil Meter Connected as a Milliammeter and Ammeter This meter can measure currents of various values, but most of the current must be shunted around the meter.

Shunt: To parallel one circuit with another.

The shunt acts just like a highway bypass around the urban area of a large city. Most of the passenger traffic and all the truck traffic is routed (or shunted) around the downtown area, while only a small fraction of the total incoming traffic goes through the city. Similarly, the more current we measure, the more current goes through the shunt because the meter can never carry more than 1 mil of current by itself. For example, when the meter reads 250 mils of current, 249 mils must pass through the shunt and 1 mil through the meter. When the meter reads 1 ampere, 999 mils must pass through the shunt and 1 mil through the meter. When the meter reads 10 amperes, 9 amperes plus 999 mils must pass through the shunt and 1 mil through the meter.

Essentially, the meter and the shunt appear to the circuit as two resistors of unlike value in parallel. The resistor of lowest value, and the one carrying the most current, is the shunt. The resistor of highest value is the meter. The meter shunt has a very low value of resistance—just a fraction of the resistance of the meter. In practice, shunts are made up of resistance wire of very low ohmmic value—for example, manganin, an iron-manganese alloy, on the low-current ranges. On the high-current ranges the shunts are frequently made up of copper wire. For example, on the 10 ampere range, the shunt may be made up of a piece of number 14 gauge copper wire, two or three inches long.

Although a formula exists for computing the resistance of the meter shunt used on various current ranges, it has little application for us since the values computed are never listed in parts catalogs or available from the shelves of an electronic supply house.

The usual practice is to determine by "cutting and trying" the shunt values for one meter of a production series and then to duplicate these prototype, or first model, values.

Measuring Heavy Currents (AC)

It becomes impractical to measure very heavy currents with a conventional meter because of the resistance introduced by the test leads, the pin jacks, and the selector switch, if any. The Amprobe meter shown in Figure 6.5 makes it possible to measure heavy currents without interrupting the circuit. Pressing on the plastic lever shown at the upper left opens the soft iron jaws above. The jaws then close around the current-carrying conductor, which acts as the primary winding of an isolation transformer. The secondary voltage is then rectified to DC by a diode, and the primary (conductor) current can be read directly on a suitably calibrated meter scale. With the addition of a battery to provide a DC source and of suitable multiplier resistors, the Amprobe meter can double as an ohmmeter or a DC voltmeter.

Figure 6.5 Amprobe Model RS-3 Test Meter Capable of measuring either heavy currents or resistance, this meter is particularly valuable in appliance servicing. (Photo courtesy of Amprobe Instruments, a division of Core Industries, Inc.)

Connecting Voltmeters and Ammeters

Voltmeters and ammeters must be properly connected to ensure that the circuit works properly and to avoid serious damage to the meters themselves. Referring to Figure 6.6, note that in this circuit the voltmeter is always connected in parallel with the source and the load. Connected this way, it can give a true measure of the electrical pressure that the source is applying across the load. If the voltmeter were accidentally connected in series with the circuit under measure, its high internal resistance might either prevent the circuit from operating or interfere with its proper operation.

An ammeter is always connected in series with the source and the load so that all current flowing

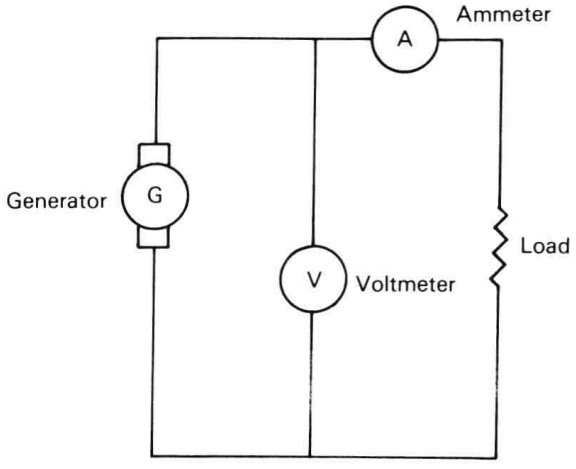


Figure 6.6 Permanent Connection of a Voltmeter and an Ammeter to Measure Power in a Typical Circuit The voltmeter is always connected in parallel with the source and the load. The ammeter is always connected in series.

in the circuit passes through its terminals. In testing a particular circuit for the value of current flow, it is always necessary to first open the circuit—for example, by unsoldering a wire from a terminal—and then to insert the ammeter in the open section so that its moving coil and shunt complete the circuit. When the test has been completed, it is necessary to disconnect the meter and solder the circuit together again.

An ammeter, because of its extremely low internal resistance, is always very vulnerable to accidental connection of its terminals across a source of voltage. If it were to be accidentally connected across the circuit under measure—the normal position for a voltmeter—a very heavy current would flow through the ammeter and probably destroy it.

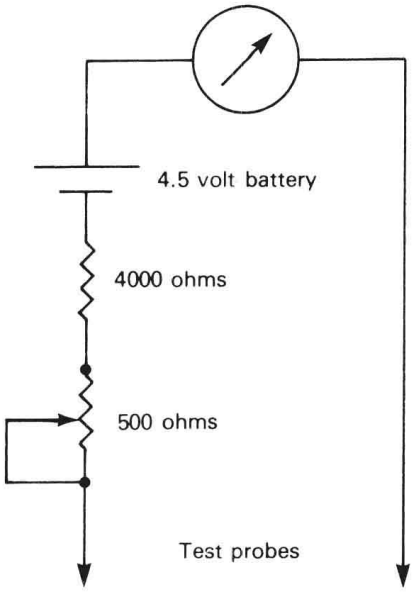


Figure 6.7 A 1 Mil Meter Connected as an Ohmmeter Designed to measure electrical resistance, the ohmmeter has a very simple circuit.

Measuring Resistance and Continuity

Ohmmeter: A meter designed to measure electrical resistance.

In Figure 6.7, we see the 1 mil D'Arsonval meter connected as an **ohmmeter.** The ohmmeter's circuit is actually very simple. In series with one lead going to the meter movement is a battery, a limiting resistor, a variable resistor, and finally a test probe. The other side of the circuit consists merely of a test lead attached to the meter terminal and terminating in a test probe.

It is convenient to use three 1 1/2 volt flashlight batteries with a 1 mil meter movement as a voltage source. The 4 1/2 volts these cells provide determine the total value of the fixed and variable resistors that are actually connected in series. The formula for computing the total value of resistance is as follows:

$$R_{total} = \text{battery voltage/full-scale meter current}$$

Substituting the values we have mentioned, we can solve the equation as follows:

$$R_{total} = 4.5/0.001 = 4500 \text{ ohms}$$

Usual practice is to make most of this resistance fixed and only part of it variable. The variable resistor is used to compensate for changes in the battery voltage as the battery ages. In this case, 4000 ohms would be a good figure to use for the fixed resistance, which dictates 500 ohms for the variable resistor (a readily available stock value).

Making Resistance Readings

With all the components connected, we can now touch the test probes together. The meter should immediately indicate a current value near full scale. Now vary the variable resistor's knob slightly until the meter hand falls over the last scale reading to the right. This operation is called *zeroing the ohmme-*

ter. It is now ready to measure an unknown resistor. Place this resistor between the two test leads. Since it is in series with the meter circuit, the meter current will drop, and a reading of less than full scale will be obtained. Low values of resistance will result in scale readings just a little less than full scale. High values will result in meter readings well toward the left of the calibrated scale since they greatly reduce the meter current. If you have an assortment of fixed-value carbon resistors, you can use them to calibrate the meter scale before using it on unknown resistors.

Checking Continuity

Continuity tests are checks to see if a particular circuit is continuous (unbroken). Usually in such tests the exact amount of resistance measured is of little consequence. For example, a continuity test is frequently applied to the primary and secondary windings of radio frequency or audio frequency transformers to determine the cause of a malfunctioning or inoperative circuit. When the test probes are connected across the ends of the various windings, current will flow, and the meter hand will move across the dial scale if the windings are continuous. If the winding is *open circuited*—that is, if it has a break in it—the meter hand will not move. Normally these transformers will operate satisfactorily if there is no break in their windings. They will be inoperative if one winding is open circuited. Therefore, continuity tests are very helpful in troubleshooting electronic equipment.

VOLT-OHM-MILLIAMMETERS

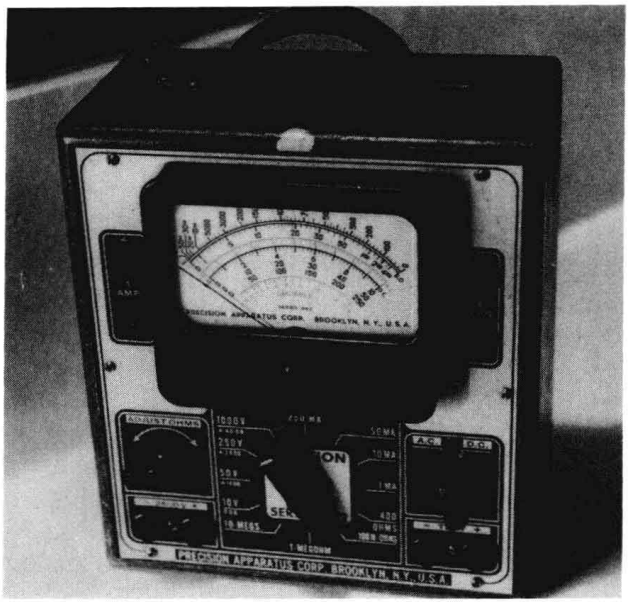
Volt-ohm-milliammeter (VOM): A test instrument that measures voltage, current, and resistance.

For over fifty years the favorite piece of test equipment of the electronics technician and experimenter has been the **volt-ohm-milliammeter.** One of the earliest multiple-function test instruments was made by the Weston Meter Company and sold in the early 1920s as a piece of test equipment for servicing early radio receivers. The instrument could not measure resistance—only limited ranges of voltage and current. It was necessary to use binding posts to select the necessary voltage and current scales because there was no selector switch.

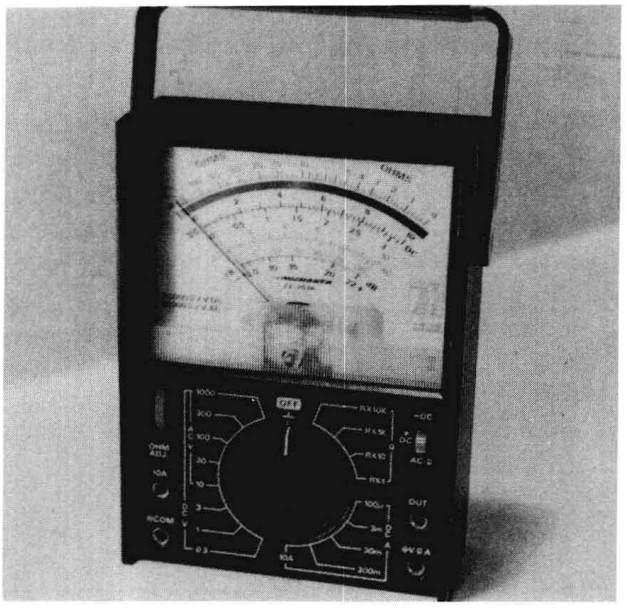
Figure 6.8A shows a somewhat later test instrument of the VOM type. Manufactured by the Precision Apparatus Company just prior to World War II, this design became very popular with radio receiver repairers and experimenters. It was easy to use, had a dial scale that was easy to read, and was ruggedly constructed. Millions of these test units were sold, and in many radio repair shops they were the most important pieces of test equipment. Many Precision VOMs are still in excellent condition today, and some electronics technicians prefer them to modern gear, particularly for working around high voltages. For comparison, Figure 6.8B shows a Micronta VOM sold in 1978. About one-third the size of the Precision meter, the newer model offers the same basic features but has more voltage, current, and resistance ranges and a more sensitive meter movement. The Micronta VOM was manufactured in Korea and distributed by Radio Shack stores. The 1983 Micronta VOM is virtually the same as the model shown in Figure 6.8B. Voltage, current, and resistance ranges are essentially identical. The appearance of the instrument has been improved through some slight changes in the panel and the dial scale. The trend in the mid-1980s is toward digital VOMs, which are much more expensive.

Reading a VOM Dial Scale

In Figure 6.9 we see a close-up of the Precision series 842 dial scale. The first scale (reading down from the top) is the principal resistance scale, with zero resistance at the extreme right and 100,000



A. This Precision series 842 volt-ohm-milliammeter (1938 or 1939) is still in regular use, particularly for high-voltage and high-current measurements.



B. This Micronta meter (1978) has the same basic functions as the Precision meter but has more voltage and current ranges and greater sensitivity.

Figure 6.8 Typical Volt-Ohm-Milliammeters

ohms at the extreme left. (In early days M was used instead of K to indicate thousand.) On the next scale, the 0–10 serves a multiple purpose: It is used for both the 10 volt and 1000 volt ranges. Remember that when you select a range with the rotary switch, it refers to the full-scale reading available on that range. For example, the 250 volt selector switch position refers to the dial scale that has 250 volts at the extreme right-hand end of the calibrated area. Therefore, 150 and 100 volts would be about at center scale, and so on. The decibel scale, showing faintly at the bottom, is used for measuring audio signal levels in audio amplifiers.

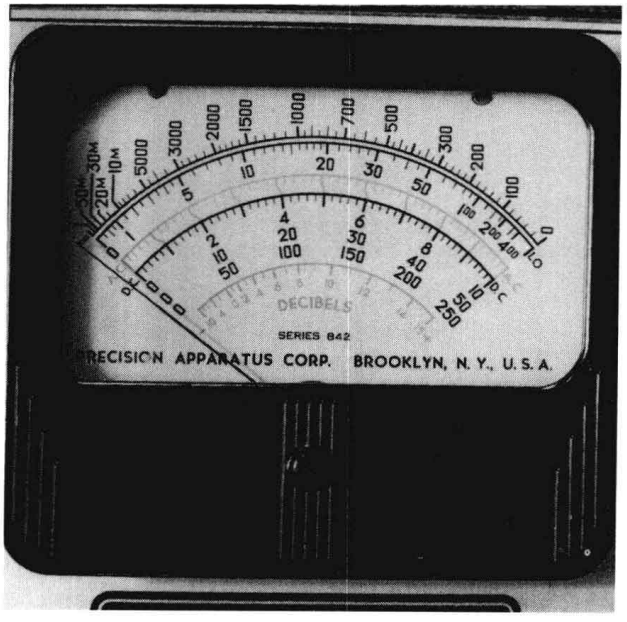


Figure 6.9 Close-up of the Precision Meter's Scale Similar scales are used on meters sold today.

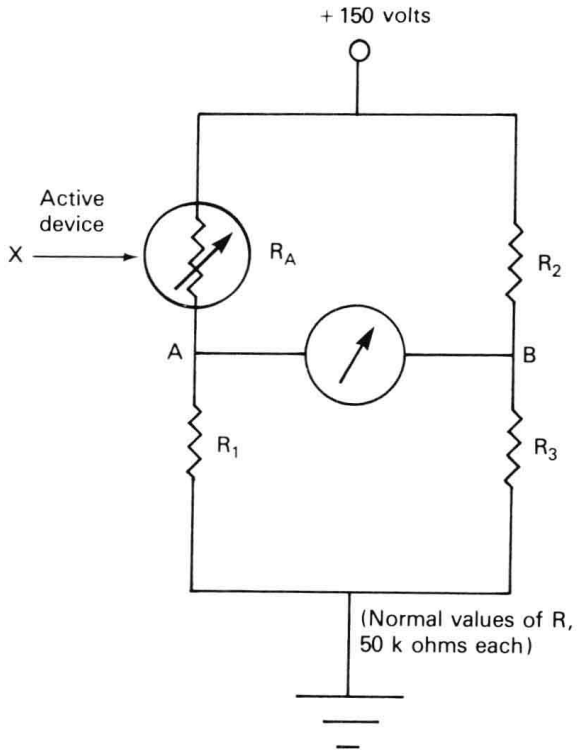
FET Voltmeters

In recent years the sensitivity and accuracy of VOMs have been improved by a new design approach that first used a vacuum tube and later a field-effect tran-

sistor (FET) to isolate the circuit under test from the current drain of the meter itself. Although we won't study transistors until the next chapter, we can understand how transistor voltmeters work by considering them as essentially variable resistors whose exact resistance varies with the voltage applied to their control device. In the case of the field-effect transistor, the control device is called the **gate.**

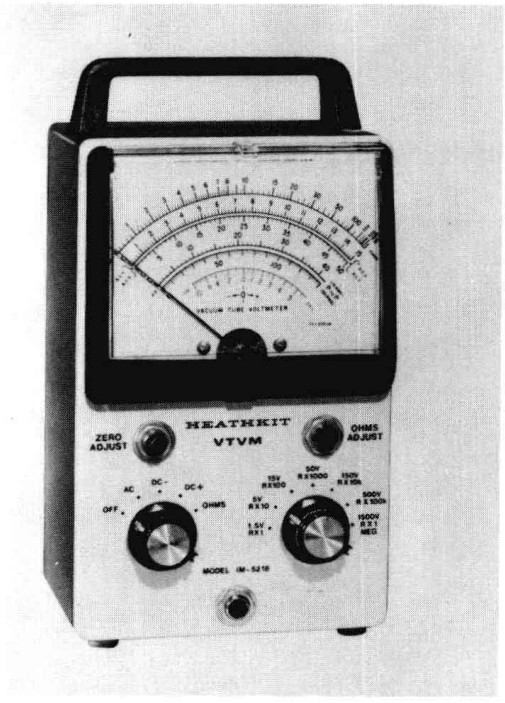
Figure 6.10 shows a circuit diagram, or schematic, of a FET voltmeter. The four identical resistors are assigned a value of 50 kilohms each, connected in series and parallel. R_A is the internal resistance of the active device (the tube or transistor). Following Kirchhoff's Law of Voltage, we know that the voltage drops across R_A and R_1 and across R_2 and R_3 are equal, and that there will be 75 volts at point A and at point B. The result is the same potential at both meter terminals and no current flow.

If the control device of the transistor is connected to a test probe and the test probe to some point in the circuit under measure, R_A will either increase or decrease in value, depending on whether the voltage at the test probe is negative or positive. The greater the voltage, the greater the resistance change will be. If the resistance of R_A becomes less, points A and B will no longer be at the same potential (the voltage at A will be higher). Current will flow



Voltage to be measured applied to point X.

Figure 6.10 Basic FET Voltmeter Circuit
Field-effect transistors have improved the accuracy of VOMs by isolating the circuit under test from the meter circuit.



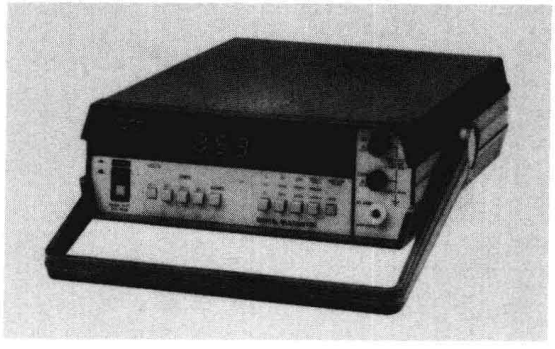
Heathkit's IM-5218 VTVM provides accurate AC and DC voltage measurements over a wide range of potentials without appreciable loading of the circuit under measure. Resistance ranges are also provided. A special probe and cable (not shown) plugs into jack at bottom of panel. (Photo courtesy of the Heath Company)

through the meter, and a scale reading will be obtained. If the resistance of R_A becomes greater, the voltage at point A will be less than that at point B, and current will flow through the meter—this time in the opposite direction. The important point about this type of VOM is that the transistor isolates the meter circuit from the circuit point being measured. Since the control element on the transistor draws no current, the circuit voltage read will be the true reading, not one flawed by the current required to operate the meter.

Digital Voltmeters

Digital voltmeters differ from other types of measuring instruments in that their display is a line of arabic numerals that expresses digitally the number of units measured. In contrast, D'Arsonval meters (in which a hand moves across a calibrated scale) and oscilloscopes (in which an electron beam traces a pattern on a phosphorescent screen covered with a calibrated grid, or *graticule*), are **analog** types.

Like its analog counterparts, the digital voltmeter employs test leads, a range selector switch, and a series of multipliers for the various ranges. Varying amounts of DC voltage from these building blocks are fed to a large-scale integrated circuit (LSIC) that performs many complex functions. First, the analog information is converted to a series of "words" made up of eight to twelve individual bits, which are processed and amplified. The bits are used to drive the display of light-emitting diodes (LEDs) to produce a numeric readout of the precise voltage encountered by the test probes. The display closely resembles that of a pocket calculator, and either LCD or LED types may be used. With slight changes in the early blocks of the instrument, it can readily read resistance in ohms and capacitance in microfarads and picofarads. Its great advantages are its extreme accuracy and the ease with which its display can be read. (Chapter 13 includes more detailed information on digital theory.) However, in such fields as AM/FM broadcasting, telecasting, audio recording, and electronic equipment maintenance and repair, analog meter types are still preferred. (An oscilloscope is an analog meter.)



Heathkit's IM-2260 digital multimeter, or VOM, provides a wide range of measurement, highly accurate digital readings, and portable operation, all at a modest price. (Photo courtesy of the Heath Company)

Reading Power in a Circuit

Refer back to Figure 6.6, which shows both a voltmeter and an ammeter permanently connected in a conventional circuit. The power being expended in this circuit can easily be determined by reading the two meters, substituting the meter readings in the equation $W = E \times I$, and then solving. However, this operation ties up two meters permanently and requires someone to take the readings and perform the calculation.

A better approach, and one widely used in industry, is to use an AC or DC wattmeter, as shown in Figure 6.11. Instead of a permanent magnet, an electromagnetic field is supplied by two coils connected in series with one side of the line. A moving coil similar to that used with the D'Arsonval meter movement is mounted in the field generated by these two coils, or field windings. The moving coil is connected between one side of the line under measure and the opposite side of the line. Since so little turning power is needed, a limiting resistor is usually connected in series with one side of the armature to limit current flow. The turning effort of the armature

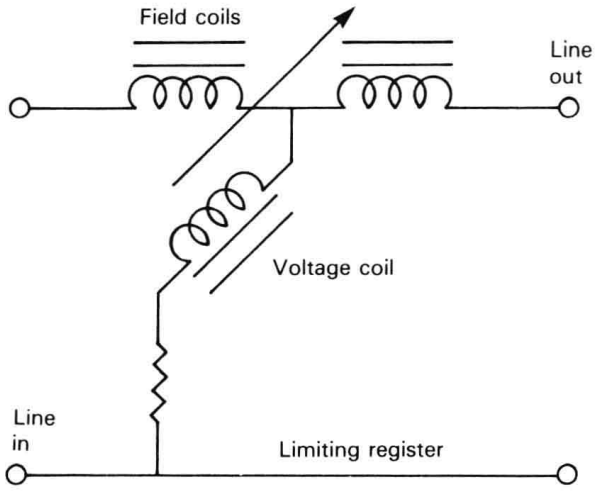


Figure 6.11 AC or DC Wattmeter The voltage coil is pivoted so it can turn.

against the tension of the helical springs that limit its movement is a function of two factors: (1) the line current flowing through the field coils and (2) the voltage across the line that determines the current through the moving coil. At any particular instant, the deflection of the meter hand corresponds to the exact amount of power being expended in the circuit ($W = E \times I$). Such a device is known as a **wattmeter** and is made in both analog form (with a moving hand sweeping across a calibrated scale) and in digital form (with numbers continuously changing).

Measuring Power at Radio Frequencies

At radio frequencies the type of wattmeter we have described is not practical since the windings offer too much inductive reactance to the radio frequency current. Instead, as shown in Figure 6.12, a solid-state diode changes the high-frequency alternating current to DC, and a DC milliammeter calibrated in watts provides the power reading.

One reason for measuring the power in a radio frequency circuit is that in many types of communications, the regulations of the Federal Communications Commission require power to be periodically monitored by the operator in charge. For example, the output of a CB transmitter may not exceed 4 watts. In the amateur service, operators may not operate with over 1000 watts input power. (Input power is the product of the plate voltage and the plate current of the final transmitter stage.) Stations in the standard broadcast service must maintain the exact output power specified in their station li-

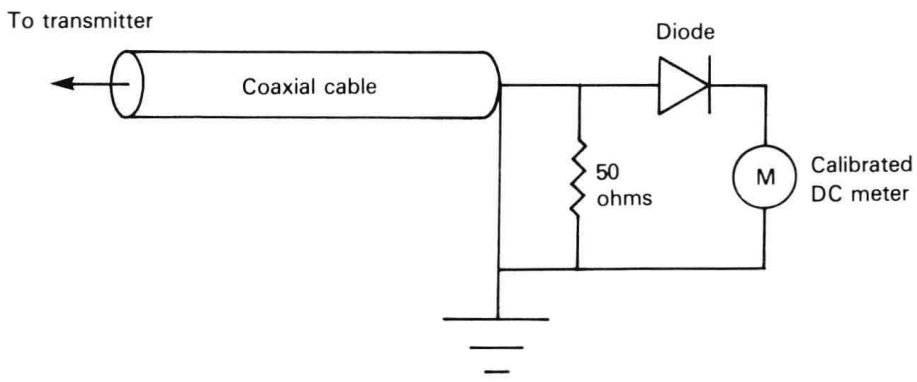


Figure 6.12 Radio Frequency Wattmeter for an Amateur or Citizens Band Transmitter Measuring AC output of an amateur or citizens band transmitter requires that the output be converted to DC by a diode and the resultant value read on a calibrated DC meter.

cense—for example, 250 watts, 10,000 watts, or even 50,000 watts.

With a suitable sensing device, the basic D'Arsonval meter is capable of a remarkably wide variety of applications: measuring the temperature of a distant star, determining the proper aperture opening and shutter speed for a camera, checking the moisture content of newly cut lumber, or indicating the received signal strength of a radio signal from halfway around the world. New applications are always being found, but the basic movement—a moving coil pivoted in the field of a permanent magnet—remains the same.

Oscilloscopes

The **oscilloscope,** closely related to a TV receiver, is basically a voltmeter with time displayed along a horizontal axis from left to right. Positive voltage is shown above this horizontal line, and negative voltage below (Figure 6.13). The trace on the screen is caused by a controlled beam of electrons within the tube striking the phosphor coating on the inside of the glass screen. When electrons strike this phosphor coating at sufficiently high velocities, light is generated and produces the trace (image). Because of the human eye's persistence of vision, the moving spot of light creates the impression of a complete waveform, not just a moving dot.

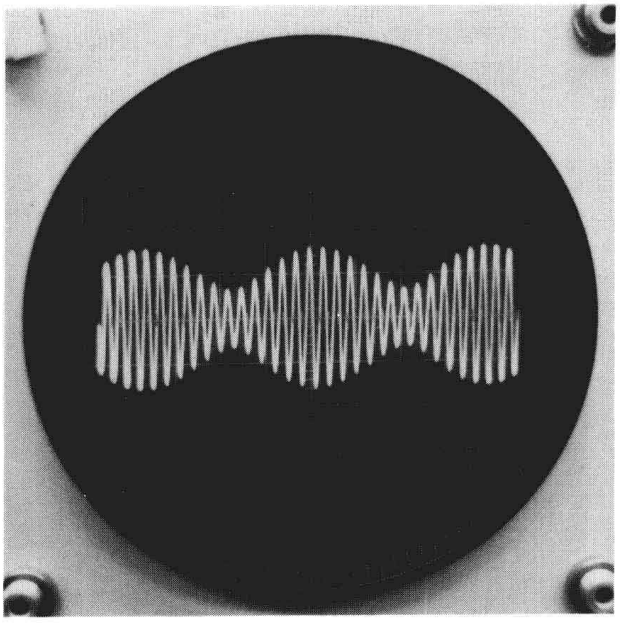


Figure 6.13 Tektronix Oscilloscope This oscilloscope displays a radio frequency wave modulated at an audio frequency rate (AM modulation). An analog voltmeter, an oscilloscope can read peak values of rapidly changing voltage with ease.

The oscilloscope has two great advantages over conventional voltmeters. First, it can actually show complex waveforms for analysis. Second, the moving beam of electrons has *zero inertia* (responds immediately) and responds to frequencies well in the megahertz. Since an oscilloscope uses a moving electron beam traveling over a calibrated scale, it is considered an analog meter. (See oscillographs, or oscilloscope patterns, in Figure 7.18.)

CELLS AND BATTERIES

Cell: A unit containing two or more electrodes (such as strips of metal) immersed in an electrolyte (a solution of an acid, base, or salt in water).

Battery: A combination of two or more cells.

Much of the experimentation that led to our current knowledge of cells and batteries was pioneered by the Italian scientist Alessandro Volta during the early 1800s. Volta was the first person to construct a **cell** and to describe and record his observations.

Volta found that when strips of copper and zinc were suspended in a strong solution of ordinary table salt (NaCl) and water, the chemical action that took place produced a voltage between the two plates that

could cause current to flow in an external circuit (Figure 6.14). As soon as the metals were lowered into the solution, the copper immediately became positive and the zinc negative. The voltage between the terminals of Volta's cell was almost exactly 1 volt, but with enough potential to perform many useful experiments. Volta soon found that by connecting two or more cells in series—that is, with the plus terminal of the first connected to the minus terminal of the second—at least twice the voltage of a single cell could be obtained but at the same current rating as that of a single cell. This combination of two or more cells we call a **battery.**

By connecting five or six cells in a series, Volta was able to conduct experiments not possible with a lower voltage. He observed that the principal factor affecting the voltage of a particular cell was the choice of metals used for the electrodes. Volta also learned that the cell's ability to supply current was principally governed by the area of the metals exposed to the electrolyte.

Today the principal cell types offered for sale are the carbon-zinc, the alkaline-manganese, the rechargeable version of the alkaline-manganese, the nickel-cadmium (known as ni-cads), the mercuric oxide, and the silver oxide. In addition are the lead-acid and Edison cells, which are never sold separately but always in combinations of cells known as batteries.

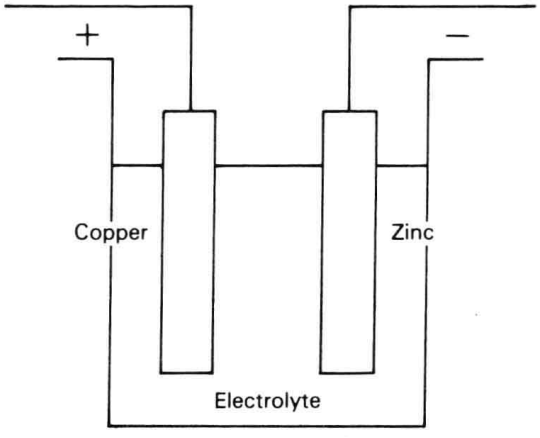
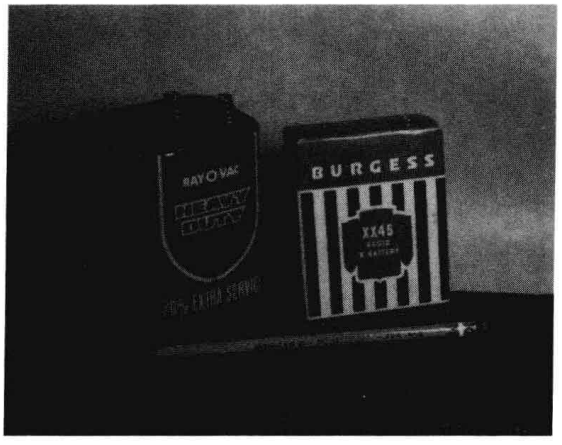


Figure 6.14 Volta's Cell The electrolyte is sodium chloride dissolved in water.

Carbon-Zinc Cells

The carbon-zinc cell, in use since the early days of the telephone, was a ready source of voltage for the earliest battery-powered radio receivers. Available in many sizes and current ratings, it provides a potential of about 1 1/2 volts over a current range from about 10 mils per hour in the case of penlight cells to about 150 mils per hour in the largest sizes. A carbon-zinc cell consists of a cylindrical zinc container that serves as the negative electrode and is usually encased in a steel jacket to prevent leakage. The electrolyte is ammonium chloride solution, which is absorbed in a sheet of cellulose material and then placed in contact with the zinc, and a carbon rod at the center of the cell that serves as the positive electrode. The area around the carbon rod is filled with a black powder, manganese dioxide, which combines with the hydrogen bubbles that form on the carbon electrode to form water. (The hydrogen bub-



Both of these batteries use zinc-carbon cells. The lantern battery (left) is made up of a number of 1.5 volt cells connected in series-parallel to provide heavy current at low voltage. The B battery (right) uses a number of 1.5 volt cells to provide high voltage at low current.

bles would otherwise prevent the electrolyte from contacting all of the area of the carbon electrode and thus reduce the cell's efficiency.)

The chemical reaction that gives rise to the difference in potential between the carbon and zinc electrodes of the carbon-zinc cell eventually converts nearly all of the zinc container to zinc chloride. Since it is impractical to reform the zinc electrode by charging the cell, the cell must be discarded when it becomes worn out, or discharged. Cells of this type are called **primary cells.**

> **Primary cell:** A cell that cannot be recharged but must be discarded when it can no longer deliver its rated voltage and current.

Carbon-zinc cells, still widely available, are being rapidly replaced in many applications by the superior **alkaline cells.** Alkaline cells, which are disposable, have much lower internal resistance than carbon-zinc cells and are therefore capable of supplying heavier currents for longer periods of time without overheating and damage. Their overall life expectancy is two or three times greater than that of carbon-zinc cells. Alkaline cells are ideal for use in transistor radios, calculators, photo-flash units, and flashlights.

Cells that can be recharged are called **secondary cells.** A 12 volt automobile battery, for example, is made up of six 2.2 or 2.3 volt secondary cells connected in series.

> **Secondary cell:** A cell that can be recharged by passing DC through it.

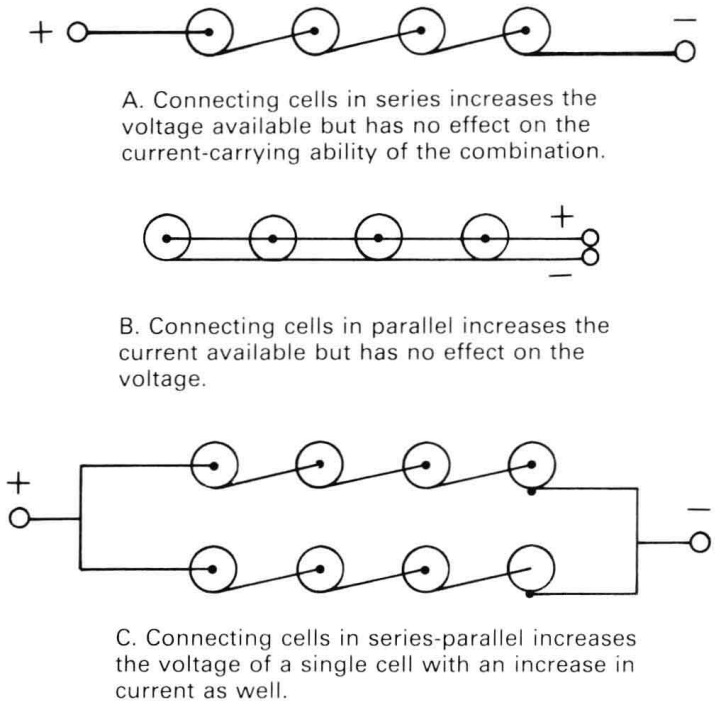
The effects of series, parallel, and series-parallel connection are illustrated in Figure 6.15. When a number of identical batteries (or cells) are connected in series, as in Part A, the voltage rating of the series is increased as additional cells are added, but the amperage remains the same as that for a single cell. When a number of identical batteries (or cells) are connected in parallel, the voltage rating of the combination is the same as for a single cell, but the amperage rating of the combination is increased as additional cells are added. In Part B, four cells rated at 1 1/2 volts and 10 amperes each are parallel connected to provide a potential of 1 1/2 volts and a current rating of 4 × 10, or 40 amperes. Notice, in Part C, that eight cells connected in series-parallel provide four times the voltage (6 volts) of a single cell at twice the amperage of a single cell, or 20 amperes.

Rechargeable Cells

In recent years alkaline-manganese and nickel-cadmium cells have been developed. They can be recharged an indefinite number of times without loss of efficiency. The alkaline-manganese cells are more costly than the carbon-zinc units but can deliver up to twice the output current and maintain their voltage levels right up to the point where they are nearly discharged. The nickel-cadmium cells are even more expensive but can provide a very heavy output current for a relatively long period of time and then can be recharged without loss of efficiency. They are used in various applications—for example, to power hedge clippers, portable electric drills, and camera and lighting equipment used by television crews for on-the-spot interviews.

Miniaturized Cells

The first miniaturized batteries were of the mercuric oxide type. They have been used for years to power such devices as hearing aids and some forms of test equipment. A more recent development, the silver oxide cell, also lends itself well to extreme miniaturization. If a camera has through-the-lens metering and an electronically controlled shutter or automatic aperture control, it is probably powered by a silver oxide cell. Digital electric watches with a numerical



A. Connecting cells in series increases the voltage available but has no effect on the current-carrying ability of the combination.

B. Connecting cells in parallel increases the current available but has no effect on the voltage.

C. Connecting cells in series-parallel increases the voltage of a single cell with an increase in current as well.

Figure 6.15 Effect of Series and Parallel Connection

time display rather than moving hands across a dial face are also powered by a silver oxide cell. Such cells, in both camera and watch applications, can be expected to last a year or more in normal service. It is a good idea to replace these cells early. During the end of their service period, they may fail suddenly without warning.

Charging Batteries

We have indicated that alkaline-manganese and nickel-cadmium cells can be recharged many times. Also, the lead-acid cells used in automotive service are constantly being charged while the car is in operation. How do you charge a battery? The first requirement is a source of DC power with a little more voltage available than the cell or battery delivers under full load. Only DC should be used. In a car, the alternator delivers about 14 volts of DC to charge a 12 volt battery. For charging alkaline-manganese and nickel-cadmium cells, you need a source that will provide about 15 to 20% more DC voltage than the normal output of these cells—usually about 1.50 volts for alkaline-manganese and about 1.25 volts for nickel-cadmium. Usual practice is to use a power transformer to step down the 110 volts from the power outlet to 2 or 2.5 volts AC. This current is then changed to pulsating DC by the use of a solid-state diode. (When connected in an AC circuit, a diode only permits electron flow in one direction to produce pulsating DC rather than AC.)

In connecting the battery charger to the batteries to be charged, it is essential to observe the proper polarity: positive to positive, negative to negative. Once the charging process is underway, check to see that the cells do not heat excessively. If they overheat, drop the charging rate by inserting a suitable series resistor in the positive lead. Complete charging data, including charging voltage and length of time required, can be obtained from the dealer where the cells were purchased.

In mountaintop microwave stations and satellites in space, the cells or batteries in the power supplies are charged by the output of suitable silicon photoelectric cells. The photoelectric cells convert the energy of the sun's rays into DC voltage and current.

In smoke alarm units for the home, only a 9 volt heavy-duty alkaline battery, such as Mallory type MN 1604, can be used for replacement. While ordinary carbon-zinc transistor batteries will fit the connectors provided, they will not operate the warning system properly. Batteries should be replaced in smoke alarm units at least once a year.

SUMMARY

One of the most useful inventions of the 19th century was D'Arsonval's moving-coil meter. The D'Arsonval movement is basically a delicately pivoted coil mounted in the field of a permanent magnet. A pointer attached to the coil moves across a calibrated scale. When current is led into the coil through helical (flat coiled) springs, an electromagnetic field springs into existence around the moving coil. The reaction between the coil's field and that of the permanent magnet produces a rotating motion that partly overcomes the opposition offered by the springs. The greater the current through the moving coil, or armature, the greater the degree of rotation.

The D'Arsonval meter, with the addition of a handful of resistors, can be used to measure a wide range of voltages and currents that the meter movement by itself would be unable to do. The resistors used to extend the voltage ranges are called multipliers and are always higher in resistance than the meter. Those used to extend the current ranges are known as shunts and are always lower in resistance than the meter movement. By the addition of a suitable battery and the proper limiting resistor, the same movement can be used to measure resistance. By switching in a suitable diode rectifier to change AC to pulsating DC, the basic meter can also be used to measure a wide range of AC voltages. A test instrument that brings all these functions together in a single case is usually known as a volt-ohm-milliammeter (VOM). The VOM, which has a hand moving across a calibrated scale, is an analog meter. Through the use of sensors, units that turn one kind of energy into another, the basic D'Arsonval meter can measure the light from a distant star or the noise level created by passing traffic.

In recent years a new type of meter, the digital meter, has been developed. It measures voltage, for example, in the form of a line of digits, or numbers. For some types of measurement the digital meter is superior to the analog meter. But for many applications, the conventional analog D'Arsonval meter is still preferred and, therefore, will be manufactured and sold for many years to come. The oscilloscope, in which a moving beam of electrons traces wave patterns on a phosphor-coated screen, is another type of voltmeter in wide use. It is classified as an analog type.

Knowledge of cells and batteries extends well back into the 18th century. Alessandro Volta, for whom the volt was named, was one of the principal early experimenters in the field. Basically, a cell consists of a container holding an electrolyte and two electrodes of different metals or different forms of the same metal. A battery consists of two or more cells connected either in series or in parallel. Connecting cells in series increases the voltage but has little effect on the current. Connecting cells in parallel increases the current-carrying capacity of the combination but has no effect on the voltage. Cells that must be discarded when run down are known as primary cells. Cells that can be recharged, or reformed, are called secondary cells.

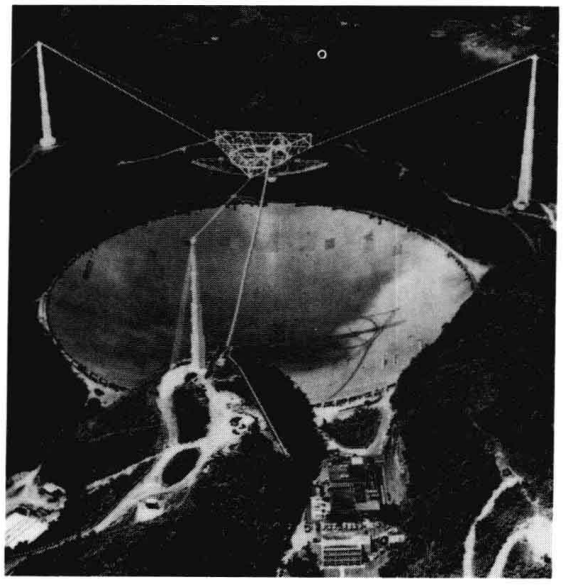
KEY WORDS AND CONCEPTS

alkaline cell
analog meter
battery
cell
continuity test
D'Arsonval movement
digital voltmeter
galvanometer
gate
multiplier resistor
ohmmeter
oscilloscope
primary cell
schematic
secondary cell
shunt
volt-ohm-milliammeter

SELF-TEST QUESTIONS

1. What was D'Arsonval's principal contribution to the technology of electronic measurements?
2. How does the dial scale of a galvanometer differ from that of a conventional voltmeter or ammeter?
3. What advantages result from moving the zero point on a meter scale to the extreme left?
4. A certain milliammeter is labeled by its manufacturer as having a "1 mil movement." In your own words, what does the label mean?
5. What is the function of multiplier resistors used in a multirange voltmeter?
6. In using a D'Arsonval meter to measure AC voltages, why is it necessary to use a solid-state diode in series with one test lead?
7. What is the function of meter shunts used in a multirange ammeter or milliammeter?
8. When a D'Arsonval meter is used to measure resistance or continuity, why is a battery needed?
9. What is the principal advantage of the FET voltmeter?
10. What is a meter called that is specifically designed to measure electrical power?
11. What is the basic difference between electrical meters of the analog and digital types?
12. How does a primary cell differ from a secondary cell?
13. When a number of cells are connected in series, what is the effect on the amount of voltage available?
14. When a number of cells are connected in parallel, what is the effect on the amount of current available?
15. Approximately how much DC voltage is required to charge a conventional 12 volt automotive battery?

7 An Introduction to Electronic Communications



The world's largest radio telescope is located at Arecibo, Puerto Rico. Made largely of wire mesh supported above a mountain valley by steel cables, it is about 1000 feet in diameter. (Photo courtesy of NASA)

Transducers
- Common Transducers
- Transducer Efficiency

Active Devices
- Edison's Contribution
- The Fleming Valve
- The Triode Vacuum Tube
- The Vacuum Tube as a Variable Resistance
- Newer Active Devices
- Comparison of Active Devices
- Active Devices as Variable Resistances
- Integrated Circuits
- Oscillation

Electronic Communications
- Historical Perspective
- The Electromagnetic Spectrum
- Propagation of Radio Waves
- Impact of Sunspots and Sunspot Cycles
- Conveying Information by Electronic Communication

Summary

Key Words and Concepts

Self-Test Questions

In this chapter we cover three important topics: transducers, active devices, and electronic communications. Transducers are already familiar to you. Commonplace devices such as your desk telephone are basically made up of two interconnected transducers—one a carbon microphone, the other a magnetic reproducer usually called a receiver. Active devices are the vacuum tubes and transistors that have made the amazing growth of the Electronics Industry possible. Electronic communications refers to how radio waves are generated and travel through space, how transmitters and receivers work, and how AM, FM, and TV modes of communication differ.

Both active devices and electronic communications are very complex subjects. However, our approach is comparatively nontechnical and low key, and the basic knowledge of electronics you have acquired in the preceding chapters should be adequate.

TRANSDUCERS

In the last decade, the word *transducer* has come into widespread use, and it represents an important concept. We can define a **transducer** very simply.

> **Transducer:** A device that takes one form of energy and transforms it into another kind of energy.

Common Transducers

Loudspeakers

Loudspeakers are devices for amplifying sound. An alternating current, surging back and forth in the voice coil attached to the speaker's cone, reacts with the fixed magnetic field. The cone then moves back and forth like the piston in an engine's cylinder and creates sound waves.

Headphones

Widely used by stereo buffs, headphones, or earphones, are made up of an electromagnet that attracts or repels an iron diaphragm when an alternating current, varying at an audio frequency rate, passes through the electromagnet's coils. The motion of the iron diaphragm causes sound waves to be set up in the surrounding air. (Later-model earphone units are constructed like miniature dynamic speakers. That is, the unit that fits over each ear is a permanent-magnet speaker with magnetic pole piece, voice coil, and core.)

Microphones

The most widely used microphone in broadcasting is the dynamic type, best described as a miniature loudspeaker operating in reverse. Instead of a true cone, it employs a diaphragm of an aluminum alloy glued to a coil that moves in a circular slot in a strong permanent magnet. When sound waves strike the diaphragm, the attached coil moves back and forth in the field of the permanent magnet. The movement causes a very weak AC current to be induced in the coil. Amplified, it provides a perfect reproduction of the original voice or music sound waves striking the diaphragm.

Indicating Devices

Two common indicating devices in modern electronic equipment are the standard D'Arsonval meter movement for measuring current, voltage, and resistance, and the center-zero type of galvanometer that is the basis for most electronic devices recording data on a chart. By substituting a form of ballpoint pen for the indicating hand, the galvanometer can be made to trace a pattern on a moving paper strip when pulsating DC is fed to the instrument's moving coil. Tracing pens are used in such differing applications as seismographs for measuring earth shocks

and electrocardiograph recorders for analyzing heart functioning.

Phonograph Pickup Cartridges

Of the many types of pickup cartridges, the best are based on magnetic and electromagnetic principles. A common approach is to provide a steady magnetic field from a permanent magnet. As the stylus moves back and forth in the record groove, it either moves a coil within this magnetic field or momentarily varies the magnetic field around a fixed coil. The AC current developed in the coil can now be amplified, and the sound wave from the speaker should be a very close replica of the original voice and the music sound waves in the recording studio. It is also possible for the coil to remain stationary and the magnet to move.

Piezoelectric Devices

Piezoelectric devices have a wide spectrum of applications in which various types of crystals, placed under mechanical stress, give rise to electrical charges on their surfaces. (In Greek, *piezo* means "to squeeze.") Common applications are the so-called crystal microphone and crystal phonograph pickup cartridge. The mechanical movement of a diaphragm in response to sound waves or of a stylus following an undulating record groove stresses a crystal of barium titanate. The weak, pulsating DC voltage collected from the surfaces of the crystal is then fed to the input circuit of an audio amplifier.

The same principle, working in reverse, is used to control the frequency of millions of transmitters, including those in the AM, FM, and TV broadcast service. A crystal of quartz (silicon dioxide) is subjected to an electrical strain in a suitable electronic circuit. The crystal vibrates mechanically at a frequency determined by its mechanical dimensions. As it vibrates, electrical impulses of the same frequency are generated. When amplified by a vacuum tube or transistor circuit, these impulses become the radio frequency carrier wave of the station.

Other Transducers

The last few years have seen widespread development of a new type of transducer that either generates or reflects light in direct proportion to an applied voltage. These transducers are principally used in digital watches and in pocket calculators (Figure 7.1). Most frequently encountered are the light-emitting diodes (LEDs), which are available in types that give off a red, yellow, or green luminance.

Of increasing importance in the case of digital watches is the use of liquid crystal displays. Such a display gives off no light, but its ability to reflect (or absorb) light varies with the electrical charge applied across the unit. Such displays are most effective in strong light, such as daylight, where LEDs frequently are washed out. More important, liquid crystal displays draw far less current than the LED type. Thus, a digital watch can operate for more than a year with a single silver oxide battery (Figure 7.2).

Transducer Efficiency

The efficiency of the various transducers varies widely. The efficiency of any single type of transducer will

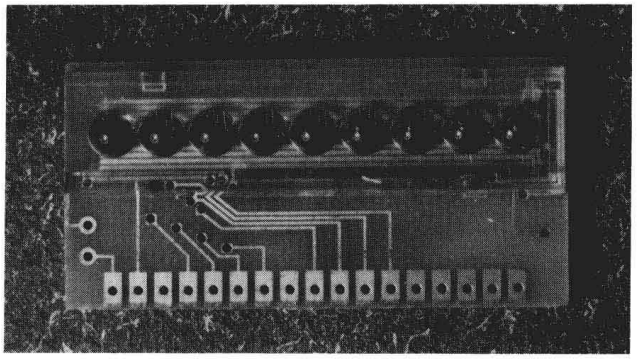


Figure 7.1 LED Assembly from a Pocket Calculator Each digit of the display is made up of seven separate diodes, as is the decimal point. By various combinations, all numbers can be reproduced.

Figure 7.2 Bulova Digital Watch with a Liquid Crystal Display Liquid crystal displays are most effective in daylight.

also vary, depending upon the manner in which it is used. For example, an ordinary 6 inch permanent-magnet loudspeaker, when simply held in the hand, shows very low efficiency in transforming electrical energy into sound waves because it is ineffective in setting any substantial amount of air in motion with the small area of its moving cone. Mounting the loudspeaker on a large square of plywood or in a reasonably large cabinet improves efficiency by perhaps 10% or 12%. (The rest of the audio energy is dissipated as heat.) Attaching the loudspeaker to a horn of considerable length and with a large opening improves efficiency even further. For example, Western Electric claimed 30% to 40% efficiency for their installations in motion picture theaters, where a special, heavy-duty loudspeaker was used as a driver unit and the horn was perhaps 25 to 30 feet long and several feet square at its opening. Some theaters still use such speaker units, although they are now considered obsolete.

ACTIVE DEVICES

Active device: An electronic component having the capability of controlling the electron flow through its structure over a very wide range of current values.

Although diodes (two-element devices) are sometimes considered active devices, in this chapter we consider vacuum tubes with three or more elements and transistors as the most important active devices. Such vacuum tubes and transistors have two unique characteristics: **amplification,** the ability to make a large signal from a small one; and **oscillation,** the generation of either an audio frequency signal (approximately 50 to 15,000 hertz) or a radio frequency signal (25 kilohertz and higher). These two characteristics account for at least 70% of the current application of vacuum tubes and transistors in the Electronics Industry. The characteristic of amplification makes oscillation possible.

Edison's Contribution

The first active device was developed in the laboratory of America's most prolific inventor, Thomas Edison. In Menlo Park, New Jersey, one of his greatest inventions, the incandescent lamp, was in trouble because the new improved filament, which should have burned without interruption for hundreds of hours, failed in 25 hours or less. But Edison had one strong clue to what the problem might be: Most of the premature failures took place at the positive

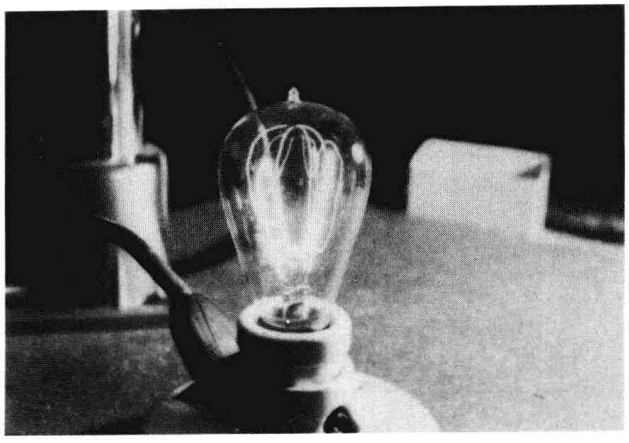


Figure 7.3 Early Edison Electric Lamp Note the looped filament structure and the sealed tip on the bulb for evacuation of gases. Filament failure in this bulb gave rise to the development of the first diode tube.

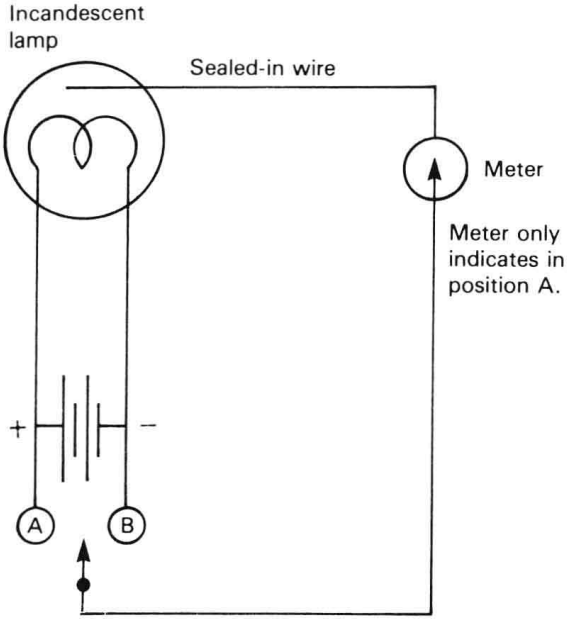


Figure 7.4 The Edison Effect When the switch is thrown to the negative position (B), electrons are repelled back to the filament in the lamp, and the meter indicates no current flow. With the switch in the positive position (A), electrons are attached to the electrode, and the meter indicates current flow.

end of the filament. He reasoned correctly that the positive end of the filament was being bombarded by negative particles that weakened it and caused it to fail prematurely. Seeking to prove his theory, he sealed an additional wire into the lamp bulb and once more created a vacuum with a mercury pump (Figures 7.3 and 7.4).

With a milliammeter in series with the added wire, he touched a lead from the other terminal of the meter to the negative terminal of the lamp. There was no indication of current flow. Then he touched the positive terminal. The meter indicated a substantial amount of current. He reasoned that negatively charged electrical particles (we would say electrons) were being boiled off the incandescent filament by the intense heat and then were being attracted to any object having the opposite electrical charge (positive). In 1884 Edison took out a patent covering his discovery. His description of his findings are contained in one of his notebooks, which you can see at the historical museum at Menlo Park.

The Fleming Valve

Although Edison had invented the first diode vacuum tube, he failed to see the importance of his finding. It remained for his chief engineer, Ambrose Fleming, to apply this invention to radio communication. By then employed by the British Marconi Company, Fleming found that the diode tube could be used as a **detector** (a device that separates a radio frequency carrier wave from the voice or music impressed on it). Marconi receivers soon appeared with the new device, which replaced the unstable coherer or crystal detector. It became known as the Fleming valve because of its ability to control the flow of electrons (Figure 7.5).

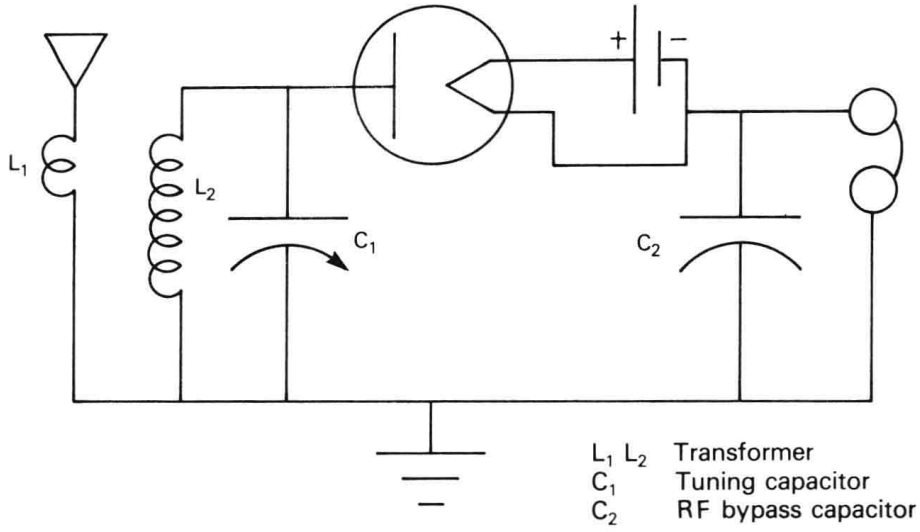


Figure 7.5 The Fleming Valve The tuned circuit L_2C_1 applies a radio frequency wave to the anode of the Fleming valve, which is actually a diode. Its filament is the cathode (negative element). When a modulated radio frequency wave is intercepted by the antenna, a pulsating DC current flows and produces sound waves of voice or music.

The Triode Vacuum Tube

The Fleming valve could only detect a radio wave. It was not capable of amplifying a weak signal into a stronger one. The problem can be illustrated by a "smoking gun" example: A very small squeeze on the trigger, the control device, will result in the release of a very large amount of energy as the cartridge is detonated. Needed was an electronic device similar to the trigger of a gun—a vacuum tube whose flow of electrons from its filament (cathode) to its plate (anode) could be varied over wide limits by some form of sensitive control device. The control device would be so sensitive that the weakest signal from a distant broadcasting station would cause a substantial change in the number of electrons moving.

In the early 1900s, a number of researchers were working on such a control device without much success. At stake was the future of long distance telephone and worldwide radio communication. The breakthrough came when Lee De Forest created the **triode** vacuum tube by adding to the diode tube a "trigger" composed of a square of wire mesh mounted directly astride the "electron highway" between the filament (the source of electrons) and the positive plate (the target for these same electrons). All the electrons flowing to the plate must first pass through the meshes of what De Forest called the *grid* (because of its fancied resemblance to the grids used on top of a wood stove for making toast). De Forest found that when he applied the weak output of a telephone's microphone (also known as a transmitter) to the tube's grid while he listened on a headphone connected in the tube's plate circuit, the signal in the plate circuit was louder than the original signal applied to the grid. For the first time, signal amplification had been obtained.

Before the triode vacuum tube, the longest telephone call that could be made from New York was to Kansas City or Denver, and connections were uncertain. Now there was no limit. Triode means a vacuum tube with three functioning elements: a filament that gives off electrons, a grid that controls

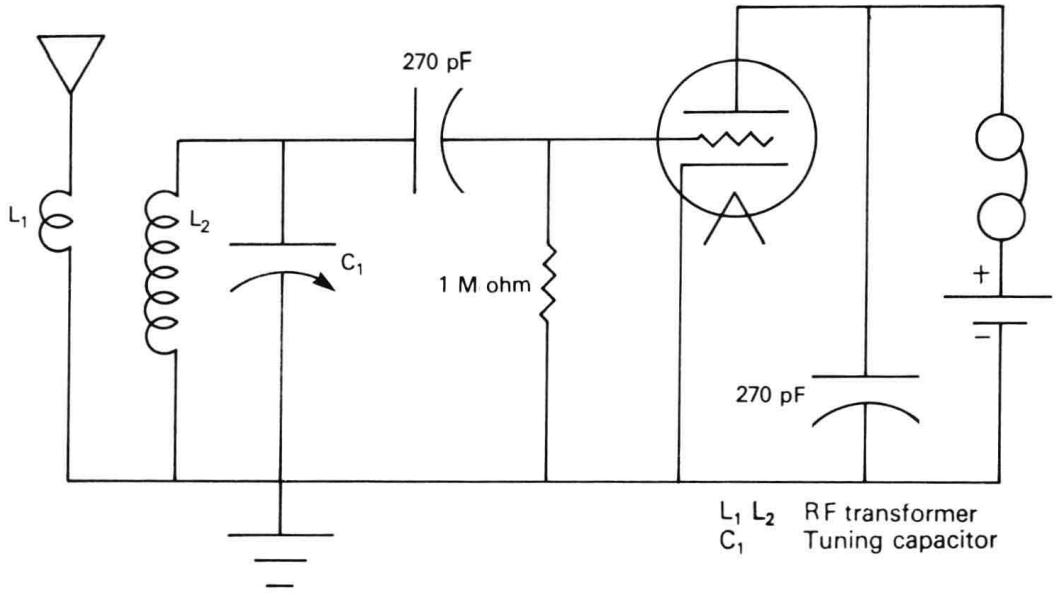


Figure 7.6 De Forest's Triode as a Grid-Leak Detector L_1 couples energy from the antenna to L_2, which is tuned to resonance by C_1. The 270 pF capacitor couples this energy to the grid of the triode, which bypasses stray radio frequency currents to ground. Headphones convert the tube's varying plate current to sound waves.

them, and a plate that attracts them with its positive charge from an external battery (Figure 7.6).

De Forest lost little time in filing a patent application for his three-element tube, which he chose to call an *audion*—probably because he could foresee immediate applications in audio frequency equipment. Learning to use the same tube at radio frequencies required some refinements in the basic design. First, the grid was changed from a square of wire mesh to a spiral of wire surrounding the filament. Second, it was necessary to remove more air from the glass bulb and make the device truly a vacuum tube. The improved triode now worked for the first time at radio frequencies and operated with improved effectiveness at audio frequencies.

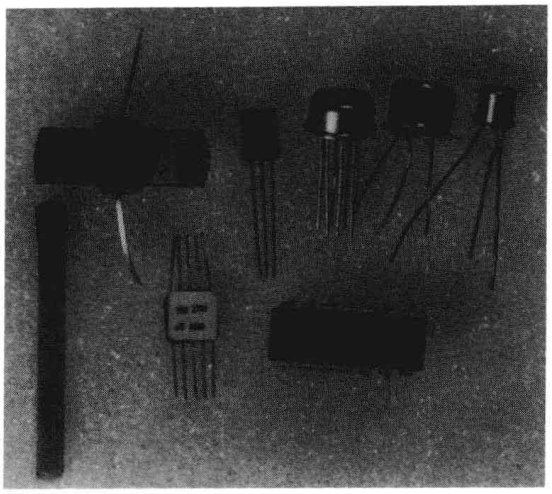
By trial and error, De Forest also learned that for optimum operation of the audion, the potential on the grid must remain weakly negative at all times. Then as the audio or radio frequency signal was applied to the grid (a signal is a varying AC or DC voltage), the grid potential would swing from negative to more negative to less negative and back again. The fixed negative potential applied at all times was called *grid bias*, and the battery that supplied it was known as the *C battery*.

The Vacuum Tube as a Variable Resistance

Over the next fifty years, engineers added two or more grid structures to De Forest's triode. The resulting tube always operated basically as a triode but with greater stability and more gain. The real breakthrough in engineers' thinking came as they began to realize that the vacuum tube was merely a highly sophisticated variable resistance—its resistance varying in response to changes in the voltage on the grid when a signal was applied.

Coming back to the smoking gun analogy, the grid corresponds to the weapon's trigger, the high-

voltage battery in the plate circuit is like the potential energy in the cartridge, and the application of the signal to the grid corresponds to the squeeze on the trigger. As the signal on the grid varies, the varying internal resistance permits more or fewer electrons to flow from the heated cathode through the turns of the grid to the plate. From the plate, the electrons pass through the load resistance to the positive terminal on the high-voltage battery, through the battery to the ground, and complete the circuit by again entering the cathode. In this way, a very small change in the voltage on the control grid produces a large change in the number of electrons flowing in the plate circuit. This effect produces amplification of the signal from grid to plate.



These transistors are of various types and have many applications. On the lower right is a typical integrated circuit, which incorporates many transistor structures, as well as diodes, resistors, and capacitors. The paper match shows the scale.

Newer Active Devices

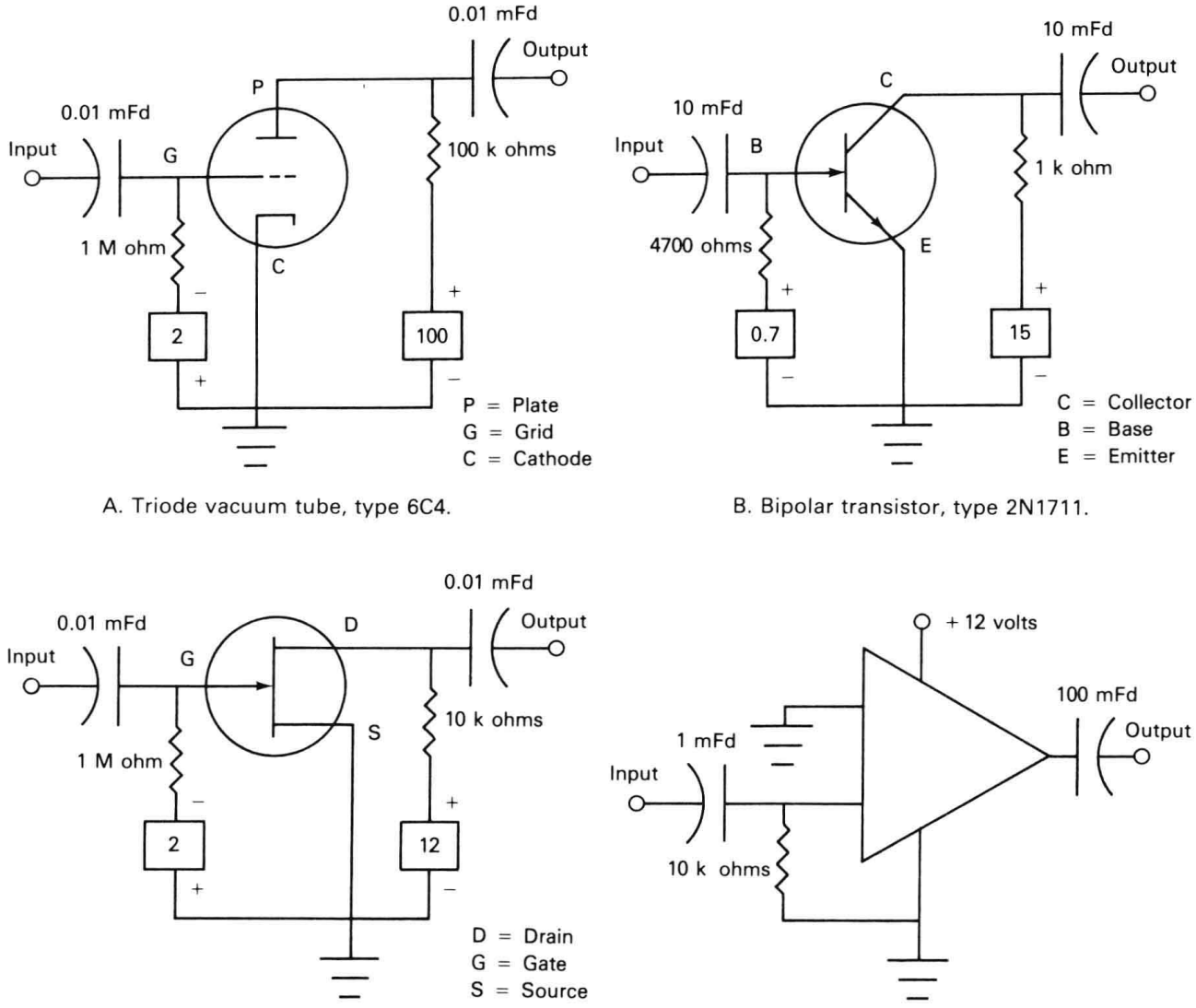
Bipolar transistor: A type of transistor with two kinds of electrical charges moving within its structure: negative charges (electrons) and positive charges (holes).

Field-effect transistor: A type of transistor that uses an electrostatic field to control the flow of electrons through the device.

The solid-state equivalents of the vacuum tube—the bipolar and field-effect transistors—also behave like sophisticated variable resistors. They permit more or fewer electrons to flow through their structure when a signal is applied to their control element. Drawing this parallel between transistors and variable resistors may be the best way for the beginner in electronics to visualize transistor operation. It is a perfectly valid approach, supported by highly trained personnel working in the industry. Remember, however, one marked difference between tubes and transistors: The electrons given off by the tube's cathode are released from the cathode's surface by the intense heating of that structure. This process is called **thermionic emission.** Not only does it require a large amount of electrical energy to produce the heating, it also requires a tube that has enough surface to radiate all of the heat generated.

On the other hand, no heating is necessary to produce free electrons within a transistor. Free electrons are available because of the manner in which the device is manufactured. When pure silicon is treated with arsenic vapor, it effectively becomes rich in electrons and is known as *N-type silicon*. If pure silicon is treated with boron in vapor form, it becomes deficient in electrons and is said to be *P-type silicon*. (N stands for negative, P for positive.)

When tiny slices of the N and P silicon are assembled in a sequence such as NPN or PNP and pressed firmly together to form a "sandwich," a transistor is created in which the comparative differences between the electron-rich and electron-deficient material create a source of electrons that will



A. Triode vacuum tube, type 6C4.

B. Bipolar transistor, type 2N1711.

C. Field-effect transistor, type EOG-133.

D. Integrated circuit, type LM-386.

Figure 7.7 Analog of Common Active Devices In this analog of common active devices, circuit components have been selected so that each active device will operate as an audio frequency amplifier.

move progressively through the device if the proper voltages are applied at the appropriate terminals. Again, no heat is required to provide a source of electrons in the transistor. The source lies in the two different types of silicon created by the vapor treatment known as *doping*.

Comparison of Active Devices

Figure 7.7 is what engineers frequently call an **analog,** an illustration designed to show both similarities and differences between various devices or

circuits. Part A shows the circuitry of a typical triode vacuum tube, a type 6C4. Part B shows the circuitry of a type 2N1711 bipolar transistor. Part C illustrates the circuit used with a type EOG-133 field-effect transistor, and Part D, a type LM-386 integrated circuit.

We will not pursue these definitions further since they are not necessary to our understanding of transistor operation at the beginning level. At this point it is necessary only to understand that all these circuits have been provided with appropriate values of voltage, resistance, and capacitance so that all will work effectively as audio frequency amplifiers. The values chosen are typical of what you might find in a commercially constructed circuit. Notice the greater simplicity of the integrated circuit shown in Part D.

Areas of Similarity

Areas of similarity exist among the first three circuits shown in Figure 7.7. Parts A, B, and C all have an input circuit designed to introduce an audio frequency signal to the structure controlling the internal resistance of the active device. In the vacuum tube, the control device is called a grid. The control device of the bipolar transistor is known as the **base,** while that of the field-effect transistor is called the **gate.**

Each circuit has a series-connected capacitor, C_1, lying between the signal source and the control element. Its purpose in all three examples is to prevent any steady DC at the signal source from affecting the highly critical bias voltage supplied by the small battery. The capacitor effectively isolates the two circuits as far as DC is concerned but readily passes the signal. Notice that the values used for the triode tube and the field-effect transistor are quite similar (about 0.01 microfarad), while a much larger capacitor is required for the bipolar transistor (about 10.0 microfarads). This bias voltage for the grid of the triode is -2 volts, exactly the same value as used on the gate of the field-effect transistor. The resistor, R_1, which prevents the signal voltage from being short circuited through the C battery to ground, is also identical for the triode and the field-effect transistor—1 megohm (1,000,000 ohms) being a common value. Note that the bias on the base of the bipolar transistor is positive—a value of about 0.6 volt being a typical value in commercial circuits.

The output circuit of all three active devices is basically the same. A capacitor, C_2, permits the signal to be applied to the input circuit of a following stage but blocks out the DC voltage from the high-voltage B battery, which would otherwise take the same path.

For the tube and the field-effect transistor, the values used for the capacitor, C_2, are remarkably similar (0.01 to 0.05 microfarad). However, for the bipolar transistor a much larger value is required (10.0 microfarads).

The term *high voltage* is always relative, but a typical vacuum tube circuit would use a power supply of at least 100 volts rating. Much of the transistor-powered equipment uses a 9 volt battery for the power source. However, in the case of complex electronic equipment, a power supply with an output in the range 12 to 15 volts provides more flexibility in circuit design. Usually this voltage is not critical, and the two transistors shown will operate well with a 9 volt battery.

The load resistor, R_2, is quite high for the vacuum tube (100 kilohms). It is very low for the bipolar transistor (1 kilohm), and intermediate for the field-effect transistor (10 kilohms).

Origin of Battery Designations

The designations A, B, and C applied to batteries go back to the earliest years in radio broadcasting. The battery that lighted the receiver's filaments, usually a 6 volt storage battery, was for no particular reason called the A battery. Similarly, the battery used to apply the high voltage to the plates of the triode tubes was known as the B battery, while the smaller battery that provided grid bias was called the C battery. This terminology has persisted, and service manuals for existing equipment often carry such notations as B+ and B−.

Active Devices as Variable Resistances

We have already discussed the operation of the vacuum tube and mentioned that the path of electron flow is from the cathode, through the turns of the grid, to the plate, through the load resistor, through the B battery, to ground, and back again to the cathode. Also, the resistance of the electron path from cathode to plate varies over a wide range in response to the signal applied to the grid. This change in resistance causes the plate current—actually the sum total of all electrons arriving at the plate from the cathode—to vary at an audio frequency rate and, in turn, to produce a voltage change at the junction of R_2. The voltage change charges and discharges capacitor C_2 at an audio frequency rate. The amplified signal is transferred to the output lead of the capacitor while blocking the passage of steady DC from the B battery.

Exactly the same sequence of events takes place with the bipolar transistor. The path of electron flow starts with the *emitter* and then goes through the *base*, the control element, to the *collector*. The resistance of this electron path (emitter to base to collector) varies widely in response to the signal applied to the base. This resistance change causes a large change in the number of electrons arriving at the collector and flowing through the load resistor to the positive terminal of the B battery. In turn, a voltage change is produced at the junction of R_2 and C_2 that charges and discharges capacitor C_2 at an audio frequency rate.

With the field-effect transistor, the series of events is almost precisely the same. The electron path within the device is from the *source* through the *gate* (control device) to the *drain*. In response to a signal applied to the gate, the resistance of the electron path (source to gate to drain) varies widely at an audio frequency rate and results in a large change in the number of electrons arriving at the drain and then flowing through the load resistor to the positive terminal of the B battery. This varying electron flow produces a varying voltage drop across R_2 that charges and discharges the output capacitor C_2 at an audio frequency rate.

All three of these active devices are sophisticated variable resistances whose internal resistance is varied in accordance with a signal applied to the control element. Even a comparative newcomer to electronics can begin to understand the operation of simple circuits employing these active devices.

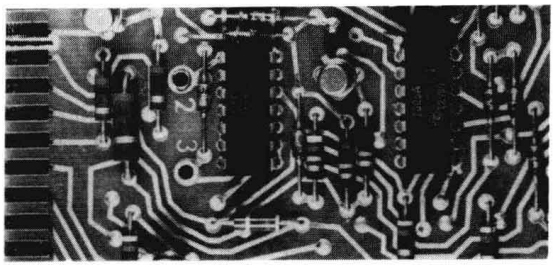
Integrated Circuits

Integrated circuit: A circuit module containing a large number of transistors, resistors, capacitors, and connecting leads.

Current reports indicate a rapid trend away from the use of separate, or discrete, bipolar and field-effect transistors and more widespread use of integrated circuits in remarkably varied applications. In the 1983 models of video cassette recorders (VCRs), personal computers, AM/FM stereo tuners, and audio tape recorders, it is frequently difficult to locate a single individual transistor. "Chips" have taken over: Regular integrated circuits (ICs), large-scale integrated circuits (LSICs), and even a few very large-scale integrated circuits (VLSICs) are widely used in practically every item of electronic equipment offered for sale.

Advantages of ICs

ICs were initially difficult and expensive to produce and were often subject to unpredictable failure. The newer ICs, however, are being mass produced with a minimum of labor. They are remarkably rugged and reliable under the most adverse conditions, and their unit price has continued to drop as the number of production rejects has decreased. Today, even a complex IC can be purchased for less than two dollars in quantities of ten. Five years ago the



A typical plug-in circuit board includes connectors, ICs, diodes, and transistors.

cost might have been twenty times greater. ICs are also lightweight and take up minimal space. Therefore, NASA's space probes and the new "smart missiles" of the armed forces use ICs almost exclusively.

How ICs Are Manufactured

The manufacturing process starts with a cylindrical crystal of chemically pure silicon that has a metallic appearance. A special saw cuts the crystal into wafer-thin discs at least 2 1/4 inches in diameter. One surface of each disc is lapped and polished to a mirrorlike finish. This side will become the IC's *substrate*, or foundation. The next step is the photochemical process. The disc is sprayed with a coating of a special plastic called a *photo-resist*. When light strikes this coating of photo-resist, the coating changes chemically where exposed (as does photographic film) and can be removed by a special solvent. Other areas of the disc's protective coating remain unchanged. Rinsed with clear water and dried, the exposed silicon is etched with a suitable chemical. Again rinsed with water and dried, the disc is ready for its next light-pattern exposure.

Production of a complex integrated circuit requires the use of a special slide projector with an ultraviolet light source for sharpness and a whole series of specially prepared slides. The first slide's exposure to the photochemical process builds the first element of an IC on a disc's substrate. The second slide's exposure provides additional circuitry, which may be followed by high-temperature baking to provide a coating of silicon dioxide on top of the new circuitry. From time to time, the exposed silicon may be subjected to an ion implantation process from elements such as boron and arsenic, which have one less or one more valence electron than silicon, which has four. This exposure produces N-type and P-type silicon, necessary in creating active devices such as diodes and transistors.

Step by step, the same pattern of operation is repeated to build up, layer upon layer, thousands of resistors, capacitors, diodes, and transistors until the IC components are complete. Finally, aluminum vapor is condensed in carefully controlled paths to provide the conductors that link the components. The chip is then complete. All the ICs on a single face of a silicon disc are entirely the same and use identical circuitry to avoid confusion. (Superficially, one IC looks pretty much the same as another, even if the circuitry is markedly different.)

When circuitry is complete, the disc is rinsed, dried, and sprayed with a clear, glass-hard protective coating. Then, like the disc shown in Figure 7.8, it is scored with a diamond scribe and the individual ICs snapped apart. Fine, gold-plated wires are attached to the terminal points on the IC either by hammering or soldering. Then, as Figure 7.9 shows, the IC is mounted in the matrix, and the gold wires are soldered to matrix pins. The matrix is then folded mechanically and filled with a hard plastic material to complete the assembly. The designation DIP for the RCA IC shown in Figure 7.10 stands for "dual in-line package" and refers to the manner in which the chip terminals are arranged in pairs.

Where an appreciable amount of inductance is required—for example, in a tuning coil designed to cover the broadcast band—current practice is to have the inductance external to the chip. Some researchers believe it will eventually be practical to synthesize inductance by various combinations of capacitors and resistors, both of which can be readily built up on the surface of the IC.

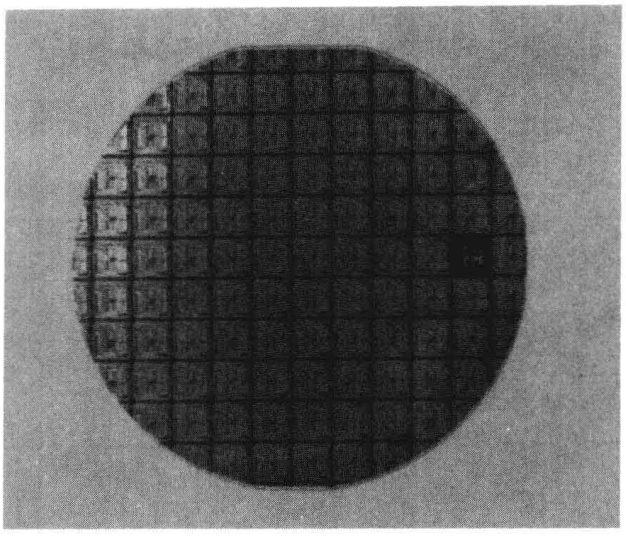


Figure 7.8 Etched Silicon Disc After photochemical treatment and ion implantation, the individual chips are ready to be separated from the disc along scored lines. The outlined chip has already been tested for proper operation.

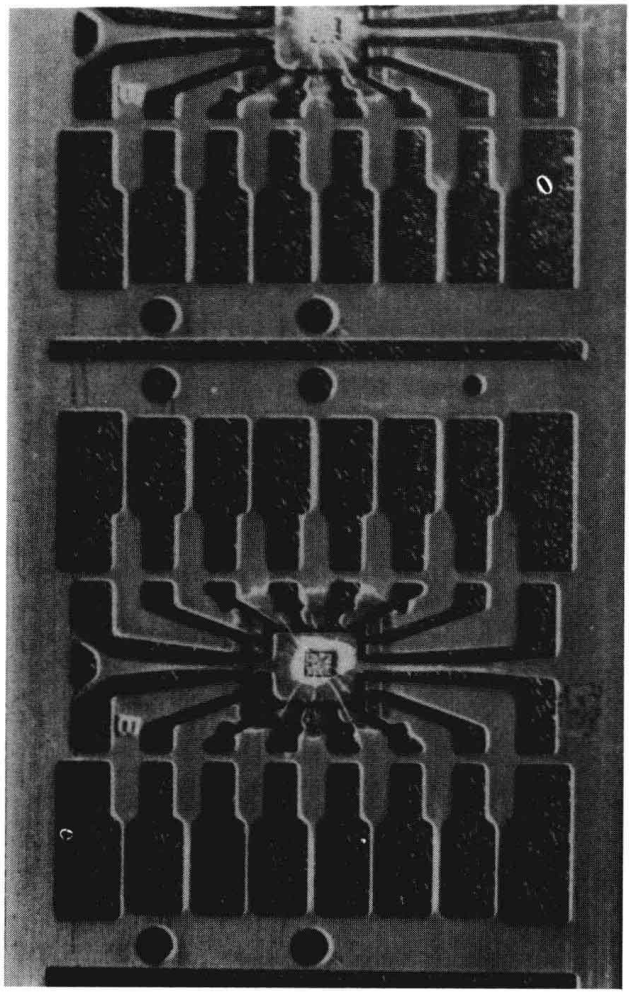


Figure 7.9 Integrated Circuit in Matrix The tiny square centered in the supporting arms of the matrix is the integrated circuit chip. It is soldered to gold-plated wires that in turn are soldered to the matrix.

ICs on Circuit Diagrams

If the surface of an IC contains thousands of transistors, capacitors, resistors, and interconnecting wires, all arranged in an extremely complex circuit, it is impractical to show all this information on a conventional schematic, or wiring diagram. Two approaches in wide commercial use provide the basic information the electronic technician must have to service equipment that makes use of many ICs in its circuitry.

The most common approach is to show the IC as an isosceles triangle with its base representing its input circuit and its apex representing the amplifier's output (see Figure 7.7D). Only the IC pins connected to the external circuit appear on the schematic (their pin numbers are shown). As ICs have grown larger and acquired more external pins, another system has gained popularity: A rectangle is used to represent the DIP package with its paired terminals. Frequently, all such terminals are numbered, but connections are shown only to those that the particular equipment employs. Unused terminals are usually wired together and either grounded or connected to a potential of +5 volts DC.

Types of IC Transistors

Many of the earlier ICs employed the bipolar type of active device in the construction process:

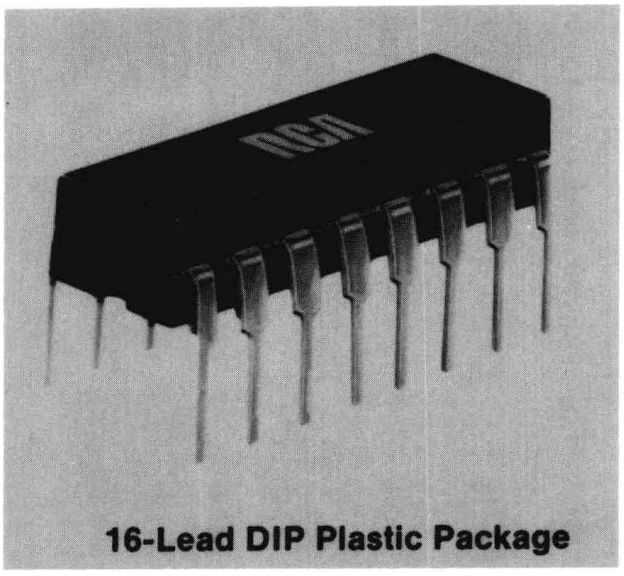


Figure 7.10 RCA Integrated Circuit This fairly complex IC is safely encapsulated in a housing of hard plastic. DIP is an acronym for "dual in-line package" and refers to the pin arrangement. (Photo courtesy of RCA)

photochemical etching, oxidation, ion implantation, and then repetition of these operations. Experimentation showed that fabrication of active devices as field-effect transistors not only offered real economies in manufacture but also better operation of the completed integrated circuit. Consequently, most of the ICs offered for sale today employ the FET type of active device in their construction.

Recent Developments in IC Manufacturing

In March 1983, a number of prominent manufacturers of ICs, including Western Electric and Intel, announced their intention to change production lines to permit the use of 6 inch silicon discs in future IC manufacture. Manufacturers believed such a move would improve the efficiency of the production process and reduce its cost. Apparently, the use of 6 inch silicon discs will soon become an industrywide practice in chip manufacturing.

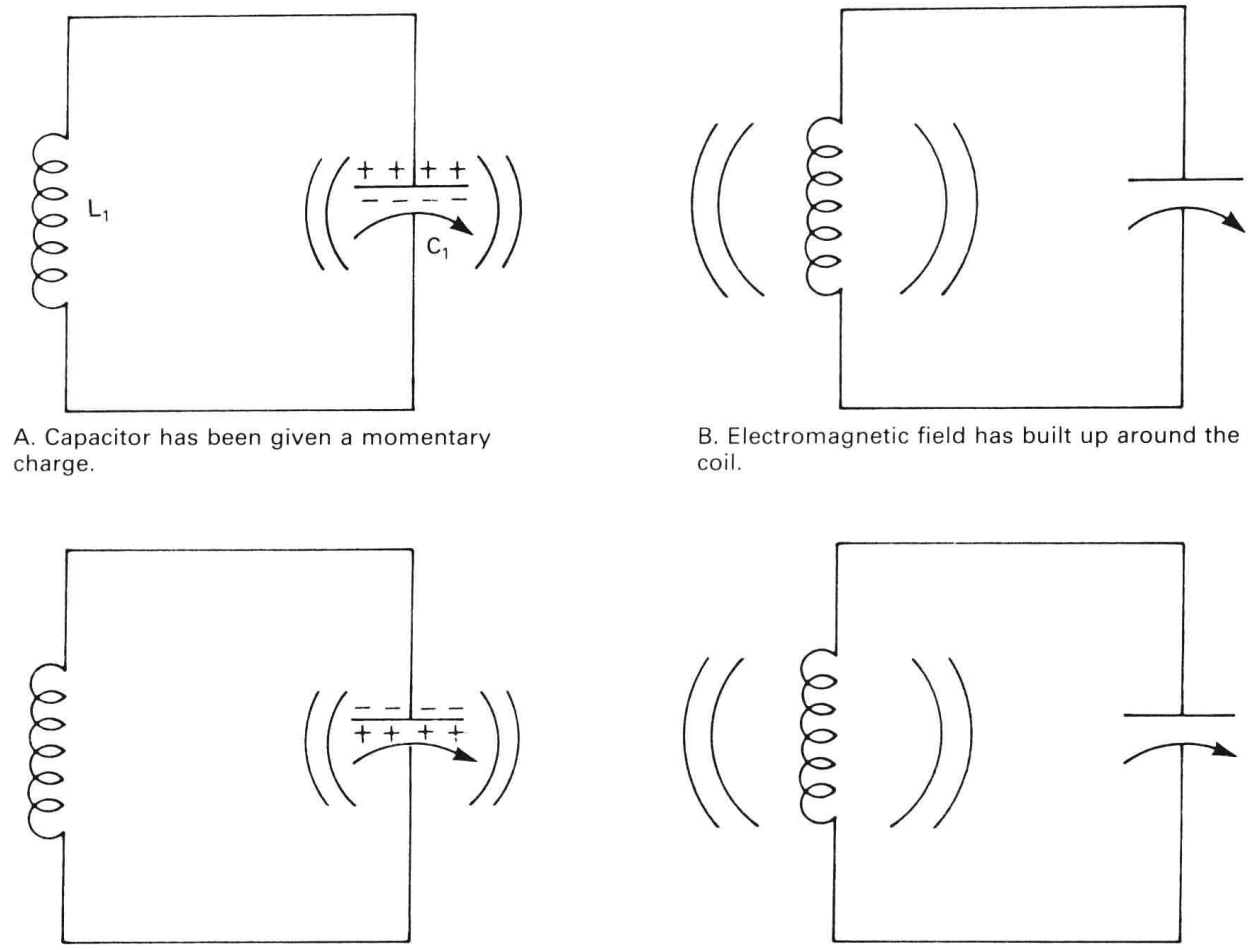
Oscillation

Oscillation: The generation of an alternating current by the exchange of energy between an inductance and a capacitance, connected together as a resonant circuit.

In a previous chapter we discussed the use of a capacitor and an inductance connected in parallel to perform the function of tuning in both receivers and transmitters. Engineers refer to such a device either as a *tuned circuit* or a *parallel resonant circuit*. A tuned circuit, when used in conjunction with an active device such as a vacuum tube or a transistor, has another important application: generating a radio frequency or an audio frequency wave. The tuned circuit plus the tube or transistor together constitute what is known as an *oscillator*, or an oscillatory circuit.

Basic Oscillatory Circuit

Referring to Figure 7.11, Part A shows an ordinary parallel resonant circuit (tuned circuit) composed of an inductor, L_1, and a variable capacitor, C_1. In Part B the capacitor is given a momentary charge from a battery, and it stores this energy momentarily in the form of an electrostatic field surrounding its plates. Then the field collapses and returns the electrical energy to the circuit. As this action sends a surge of current through the inductor (coil), an electromagnetic field is created briefly around the coil. Then the electromagnetic field collapses and returns its energy to the circuit, this time charging the capacitor to the opposite polarity (Part C). The electrostatic field around the capacitor again collapses and once more builds up a magnetic field



A. Capacitor has been given a momentary charge.

B. Electromagnetic field has built up around the coil.

C. Capacitor has been charged in opposite direction.

D. Electromagnetic field is again built up around the coil.

Figure 7.11 Basic Tuned Circuit

around the coil (Part D). If we touch the test probe of an oscilloscope to the circuit, we can see a *sine wave*, the same characteristic waveform we observed in our study of alternating current.

Unfortunately, this process of storing energy alternately in the capacitor and in the coil will not continue indefinitely because of what we call the losses in the circuit. For example, the windings of the coil have some resistance. Also, the tuning capacitor has some unavoidable resistance between the plates mounted on the rotating shaft and the wiping brushes on the frame. To keep the oscillation going, we must introduce enough energy into the tuned circuit from an external source to make up for the energy losses in the oscillatory circuit. Therefore, we always use a vacuum tube or a transistor in conjunction with the coil and the capacitor. In a suitable circuit, the active device can provide a pulse of energy at the precise time to keep the circuit oscillating (generating an AC waveform) indefinitely.

Figure 7.12 shows a transistor in a circuit suitable for producing continuous oscillation. The cir-

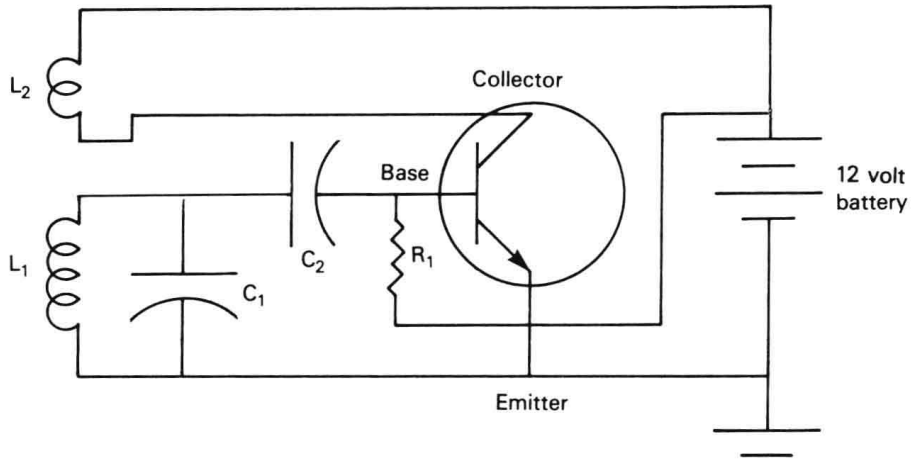


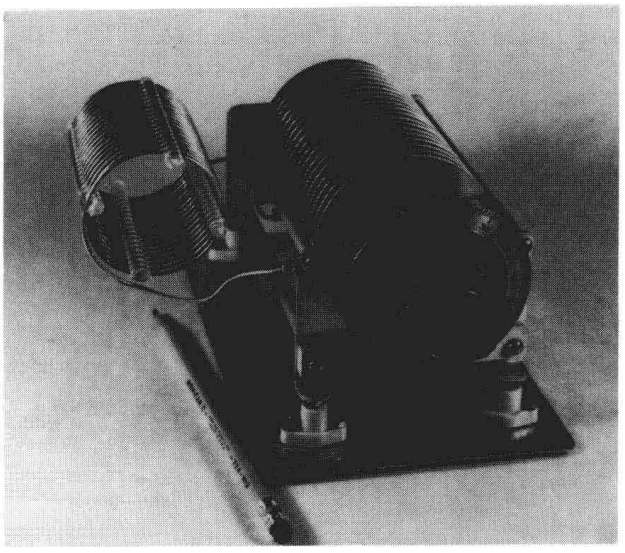
Figure 7.12 Oscillator Circuit The addition of a bipolar transistor allows production of continuous oscillation.

cuit's most important feature is the addition of a second coil, L_2, connected in the transistor's collector circuit and inductively coupled to L_1. Because of the signal amplification provided by the transistor, the signal in the collector circuit is much larger than that applied to the base through coupling capacitor C_2. L_2 usually has only about one-fourth to one-fifth as many turns as L_1, but this number is sufficient to couple enough energy to L_1C_1 to sustain oscillation. The function of R_1 is to provide a bias voltage on the base of the transistor. In practice it might be in the range from 27 kilohms to 47 kilohms. A good value for the capacitor C_2 would be 270 picofarads. At this size, it can charge and discharge rapidly enough for operation at the high radio frequencies, yet still work effectively at the low radio frequencies.

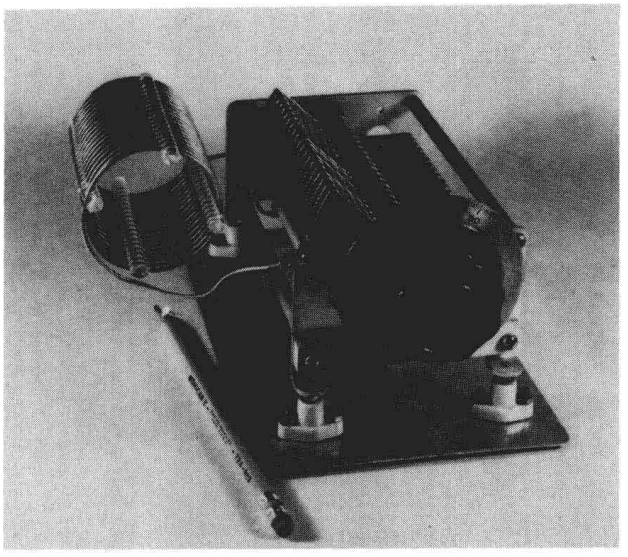
The frequency of oscillation can be varied over wide limits by merely turning the knob on the shaft of the tuning capacitor. As shown in Figure 7.13A, with the plates out of mesh, the capacitor has less capacity and the tuned circuit is in tune for a higher radio frequency. As shown in Figure 7.13B, with the plates fully in mesh, the capacitance is increased and the tuned circuit is in tune for a lower frequency. Oscillation can be made to take place on a wide range of frequencies by merely varying the capacity of the tuning capacitor—the usual approach when oscillator circuits are designed for radio receivers and radio transmitters. However, the reverse is also a possibility. If a high-quality fixed mica capacitor is used to replace C_1, tuning can be accomplished by making L_1 a variable inductor. (Move a ferrite—powdered iron—core in and out of the coil by suitable mechanical means.) The tuned circuit will then be in tune for the highest frequency when the iron core is completely withdrawn and will be in tune for the lowest frequency when the iron core is completely inserted. In a typical compact transistor radio, two tuned circuits use a fixed coil and a variable capacitor, while four tuned circuits use a fixed capacitor and a variable inductor (coil).

Verifying Oscillation

Occasionally it is necessary to verify that the oscillator is actually oscillating and generating a radio frequency wave. If all operating voltages are normal on the active device, oscillation is probably occurring. However, as a positive check, listen on a nearby radio receiver with the same tuning range. You should be able to hear a steady, unmodulated (no voice or music) carrier wave that will change its frequency when the tuning capacitor is varied by a turn of its control knob.



A. This tuning capacitor is out of mesh and provides the lowest possible capacity. Under this condition, the tuned circuit is tuned to a high radio frequency.



B. This tuning capacitor is nearly at maximum capacity (plates are nearly in mesh) and is in tune for a low radio frequency.

Figure 7.13 Tuning Capacitors

ELECTRONIC COMMUNICATIONS

The purpose of this section is to provide an understanding of how electronic communications systems operate—that is, how radio signals are generated and propagated through space, how they are affected by the earth's ionosphere, and how they are received. More detailed explanations are found in Chapters 10 and 14.

Historical Perspective

For billions of years, the earth has experienced thunderstorms, and every lightning flash has caused radio waves of widely varying intensity and wavelength. Not until the late 1800s did anyone realize that what was needed was a radio receiver and someone to operate it.

It has been known for over a hundred years that an electric arc of any kind generates radio waves effectively. Heinrich Hertz used a spark coil, much like the ignition coil in a car, in his early experiments with radio waves. The coil generated a hot spark between two electrodes, and the spark set up radio waves of about 1 meter wavelength that Hertz could actually measure. Upon Hertz's death, the same line of experimentation was followed by Guglielmo Marconi, who had superior equipment and a staff to assist him. When, in 1901, he became the first person to send a radio signal across the Atlantic (the three dits of the letter *s* in Morse code), he was using a powerful arc transmitter located at Cornwall, England, and a rather sophisticated receiver at St. Johns, Newfoundland. The receiver was equipped with several tuned circuits for maximum response to the desired signal, and both transmitter and receiver were connected to towering antennas that Marconi had found essential for maximum effectiveness. An elaborate ground system consisting of buried wires connected and tied to the transmitting and receiving equipment improved the efficiency of the overall system still further.

The arc transmitter created an immense amount of interference to other nearby frequencies, and as the number of transmitters and receivers in operation increased, a different system of generating radio frequency waves had to be devised. A further consideration was the fact that the arc transmitter was not well adapted to voice and music transmission.

A new step was taken by a research staff at General Electric under the direction of Ernst Alexanderson. After many setbacks, they were able to produce a satisfactory generator of radio frequency waves, appropriately called the Alexanderson alternator. It operated just like the AC alternator in a car but had many more pole pieces and armature coils. Turning at a very high speed, it could develop an AC wave of a frequency perhaps as high as 100 kilohertz. Since it was developing a relatively pure wave, free from harsh harmonics, the alternator was immediately pressed into service for voice and music transmission. A Christmas Eve broadcast in 1906 from Brant Rock, Massachusetts, with Reginald Fessenden at the controls, was widely heard throughout New England and by ships at sea.

Despite early success, the Alexanderson alternator was not an ideal generator of radio frequency waves. It was a mechanical device of great weight and complexity and was constantly threatened with disintegration from centrifugal forces caused by the very high shaft speed. A better solution appeared with De Forest's invention of the vacuum tube (audion) in 1907.

The audion's ability to oscillate as well as to amplify was discovered soon thereafter by both De Forest and Armstrong. Even today, hundreds of thousands of transmitters of various types use vacuum tube oscillators to determine their frequency of oscillation, while larger vacuum tubes build this comparatively small signal up to a level of 50,000 watts or more. Only within the past five years have transistors been used widely in such applications, and then only at low power levels.

The Electromagnetic Spectrum

All forms of electromagnetic radiation differ from each other only in terms of frequency and wavelength. Referring to Figure 7.14, we see that the lowest frequency normally encountered is the 50 or 60 hertz current generated by public utilities for electrical power distribution. At the same time, 50 hertz is usually considered the lowest audio frequency, while the upper audio frequency limit in terms of its

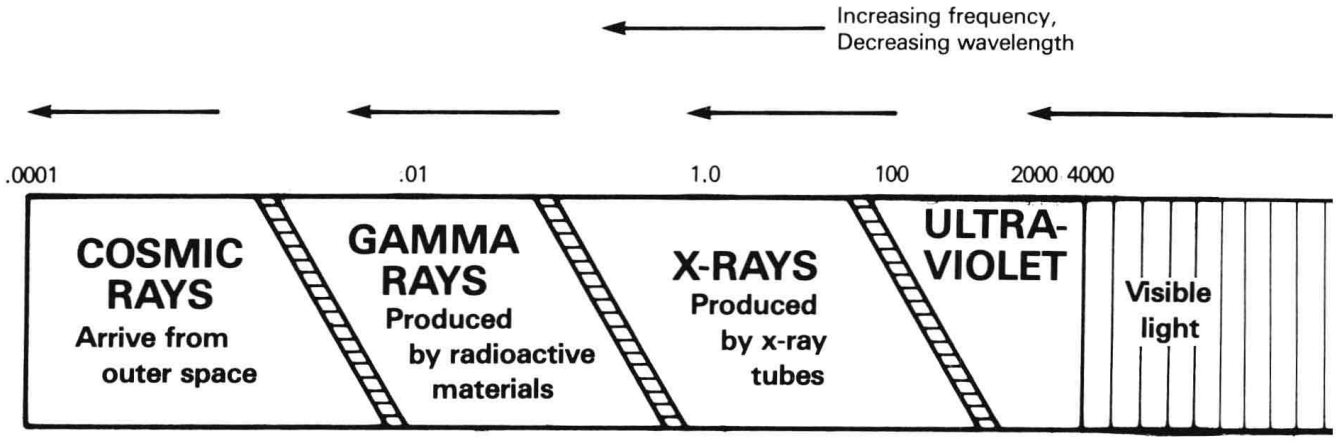


Figure 7.14 Electromagnetic Spectrum Electromagnetic waves of varying frequency and wavelength lie between cosmic rays and electric waves.

definition in most electronics references is about 20,000 hertz. Logically, frequencies immediately higher are known as *ultrasonic;* those still higher, as *radio frequency* waves. The transition from one category to another is both gradual and marked by considerable overlapping. The slanting lines in the drawing are intended to convey this concept of the gradual and overlapping transition between the major categories of electromagnetic radiation.

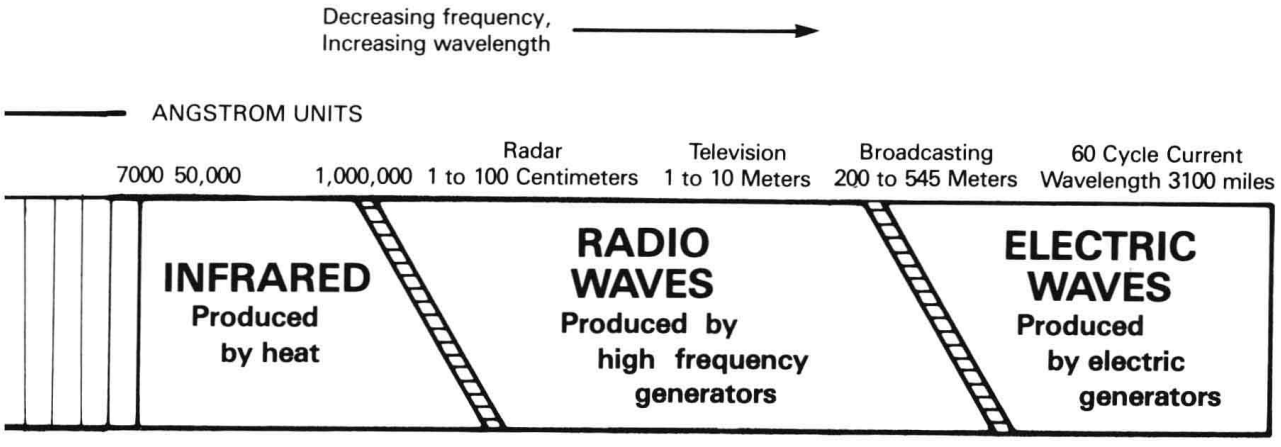
Other segments of the electromagnetic spectrum can be used for other types of communication. For example, television stations frequently send news events from the site back to the studio by modulating a beam of infrared light with the program material and then focusing it to a pencil-thin beam with a parabolic reflector. Also, public utilities commonly modulate the 60 hertz current they send over their transmission lines with voice messages and data information. It can be picked up at a distant point along the line by a specially designed receiver that recovers the voice and data information.

Propagation of Radio Waves

As far back as the Civil War, the British scientist James Clark Maxwell had mathematically predicted the existence of radio waves and had even determined their speed of propagation: 186,300 miles per second—the same as that for electromagnetic waves such as light and heat. It remained for Marconi to observe that some radio waves proceed directly to the receiving location from the transmitting antenna, while others take a more indirect path, ricocheting off the layers of ionized gases in the earth's upper atmosphere, known as the **ionosphere.**

Figure 7.15 shows radio waves taking two directions from the transmitting antenna. The portion of the wave that travels in contact with the ground is known as the **ground wave,** while that directed toward the earth's ionosphere and then reflected back to the earth's surface is known as the **sky wave.** The ionosphere is a vast layer of ionized gases of the earth's atmosphere, extending from about 50 miles above the surface outward to perhaps 200 miles or more. The energizing force in creating ions from gases such as nitrogen and oxygen is that part of the sun's radiation that lies in the ultraviolet part of the spectrum.

Figure 7.16 is a simplified sketch showing the radiation of high-angle waves returned by reflection from the ionosphere as sky waves. This illustration especially applies to those broadcast frequencies falling between 1600 kilohertz and 540 kilohertz. At



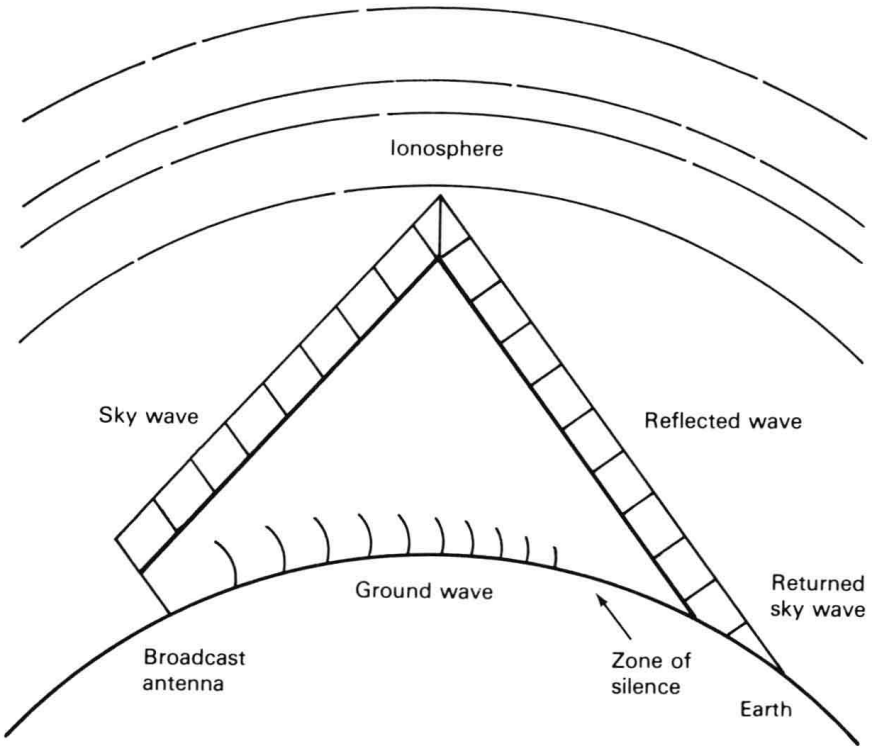


Figure 7.15 Direction of Radio Waves from a Transmitting Antenna One portion of the wave travels along the ground (ground wave), and the other is directed out toward the ionosphere and then back to the earth's surface (sky wave).

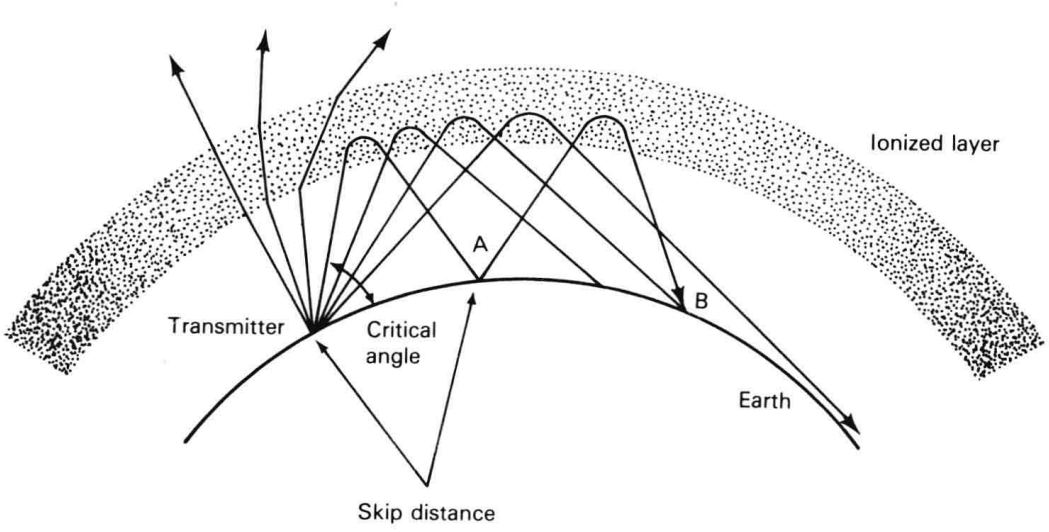


Figure 7.16 Radiation of Sky Waves Sky waves are high-angle waves returned by reflection from various layers of the ionosphere.

these frequencies there is very little return of the sky wave to the earth during daylight hours. If you are listening to a broadcast-band signal during the daytime, you are undoubtedly listening to that station's ground wave. When darkness falls, however, the sky wave is greatly strengthened as the ionosphere's reflectance for these frequencies is enhanced by the absence of sunlight. If during the evening hours you're listening to a standard broadcast station several hundred or several thousand miles away, you're unquestionably receiving the signal by sky wave, because the station's ground wave is seldom effective over a distance of 100 miles.

Figure 7.16 is more applicable to the VHFs (very high frequencies)—those lying between 3 and 30 megahertz, where the ground wave is of little consequence and communication is largely dependent upon the sky wave. Note that when radio waves strike the ionosphere at an angle approaching 90°, all or most of the energy penetrates this layer of ionized gases, and little if any of the signal is returned to the earth's surface. As the angle is made shallower, more and more of the signal is bounced back to the earth.

Since there is little ground wave at these frequencies, the signal may not be heard from a distance just a few miles from the antenna until it returns to the earth, perhaps several hundred miles away. Broadcasters once referred to this area without reception as the *zone of silence*. In more recent years, radio amateurs and citizens band operators have called this zone the **skip distance,** the most common term in use today. Note that under favorable conditions the signal returned to the earth (at point A) may be reflected from the earth's surface, where it is again reflected or refracted (bent) by the ionosphere and returned this time (at point B) perhaps another two or three hundred miles from the source antenna. This process is usually called **multihop transmission.** Under unusually favorable conditions, the signal from even a low-power transmitter can be heard with good volume more than halfway around the world. Such reception is more likely to take place on the high frequencies, such as the amateur 10 meter band (28 to 30 megahertz), and particularly during periods of high sunspot activity. A **sunspot** is a magnetic storm on the surface of the sun. The storm appears much darker than the surrounding area, hence the name.

Impact of Sunspots and Sunspot Cycles

While sunspots do not themselves directly affect the ionosphere and electronic communication, they appear on the sun's surface in greater or fewer numbers over a cycle that takes approximately eleven years to wax and wane. During the peak of the sunspot cycle, solar activity is at a maximum, with frequent magnetic storms of great intensity, an increase in the number of charged particles radiated from the sun, and an accompanying increase in the amount of ultraviolet and X-ray radiation from the sun's surface. Radiation in that portion of the spectrum is thought to be the principal cause of ionization of the earth's atmosphere that in turn makes up the ionosphere. Figure 7.17 shows the various layers of the ionosphere.

During periods of maximum sunspot activity, communication becomes possible in the VHF bands on a worldwide basis easily as high as 30 megahertz, and occasionally higher than 50 megahertz since ionization is intense enough to reflect or refract the sky wave signal from the transmitting antennas to the distant receiving antennas. Low-frequency communication is poorer than average at this time. Even the standard broadcast band (1600 kilohertz to 540 kilohertz) is noisy and erratic, and fading at night may be severe.

In periods of low sunspot activity, lack of ionization may cause the high-frequency spectrum to go dead except for local contacts because the degree of ionization of the ionosphere is not intense enough to return signals to the earth's surface. Instead, they punch through the ionosphere and enter outer space, never to return. At the same time, the low-frequency bands are at their very best. The broadcast band, for example, performs excellently: Signals are strong, noise is minimal, and fading is not severe. The same is true of the amateur 160 meter band (1.8 to 2.0

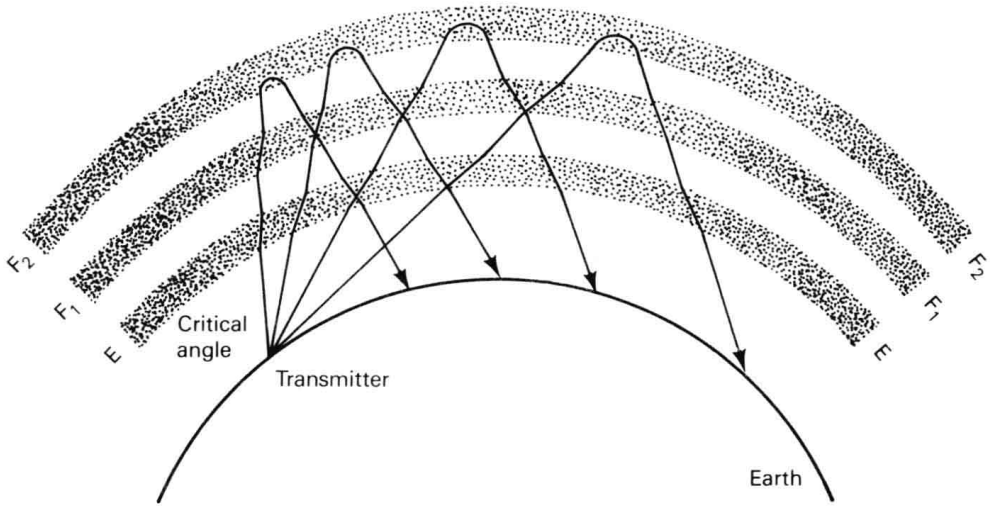


Figure 7.17 Various Layers of the Ionosphere and Their Refractive Effects The E and F_1 layers of the ionosphere bend the sky waves slightly, but only the F_2 layer bends them enough to return them to Earth.

megahertz) and, to a lesser extent, of the 75 meter amateur band (3.5 to 4.0 megahertz).

Sunspot cycle 21 peaked in December 1979. For several years prior to and after that date, the amateur 10, 15, and 20 meter bands were exceptionally good for long distance and worldwide communications. U.S. amateur stations, even those with low power and simple antenna systems, had little difficulty in working other stations in Europe, Africa, and Asia. Solar activity is now declining quite rapidly, and DX (long distance reception and transmission) opportunities are rapidly dwindling. As these bands become less effective for communication, the amateur 80 and 160 meter bands are rapidly improving, and amateurs report that signals are stronger than they have been in many years on the comparatively low frequency bands. This trend is expected to continue until cycle 21 bottoms out in about 1988 and gives way to cycle 22 with its increasing solar activity.

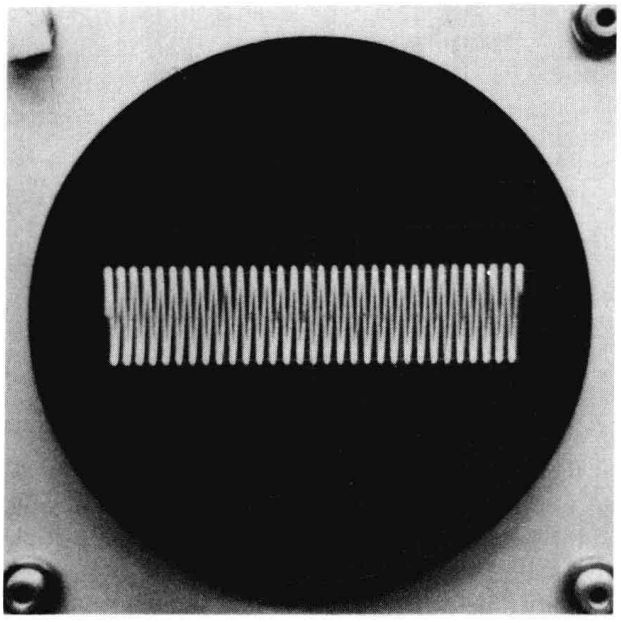
Conveying Information by Electronic Communication

Until 1906, most messages sent by wireless used a very simple technique. When the operator's telegraph key was closed, some form of electric arc generated a radio frequency wave that, in turn, was coupled to a radiating antenna. As long as the key was down, the radio wave leaving the antenna was continuous. When the key was raised, the arc was turned off, and no radio wave left the antenna. By turning the arc off and on with the hand key in accordance with the pattern of the letters and numbers making up the Morse code or the international Morse code, a distant operator equipped with a receiver and earphones could convert the dots and dashes back to letters and numbers. This technique of sending messages by radio is still used today, but the arc transmitter has been replaced by one employing vacuum tubes or transistors, and both the transmission and reception may be accomplished by electronic equipment that permits very high speed handling of messages.

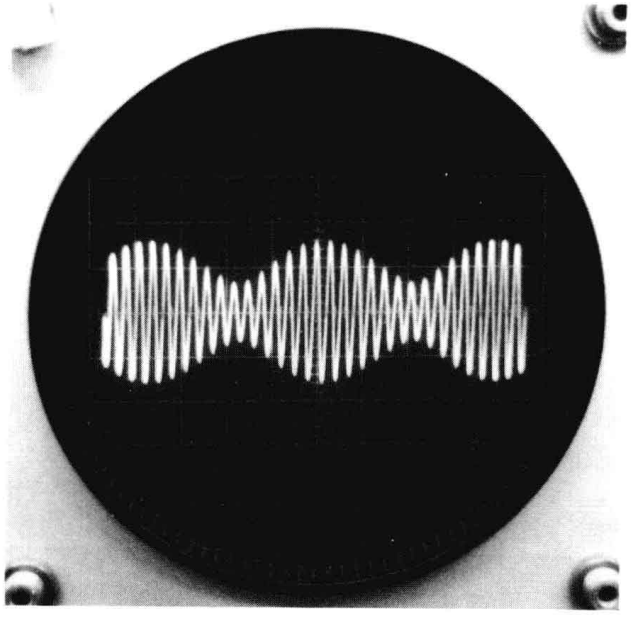
In 1906, Ernst Alexanderson and his General Electric research team succeeded for the first time in transmitting voice and music of good quality. The equipment was basically the motor-driven Alexanderson alternator, which mechanically generated the carrier wave of the broadcasting station that was continuously transmitted. (With today's transmitters, the carrier wave is generated by either vacuum tubes or transistors.) By a process known as **mod-**

ulation, voice and music were added to the carrier wave. The carrier wave's power, or **amplitude,** was varied at an audio frequency rate. At the reciever it was "detected," and the voice or music power variation, now turned to pulsating DC, actuated the metal diaphragms in the headphones to once more create sound waves. In this application, "to modulate" means to vary the power of a radio frequency wave at an audio frequency rate. This form of modulation, known as **AM,** or **amplitude modulation,** is used by all radio stations in the standard broadcast band, by most citizens band stations, and by many amateur stations.

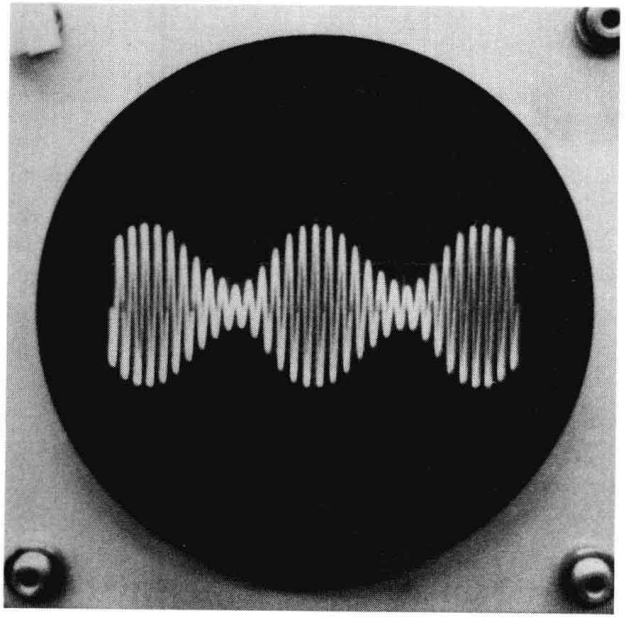
Figure 7.18A shows an unmodulated radio frequency carrier wave. In Figure 7.18B, the same carrier wave is partially modulated by a 1000 cycle audio tone. Figure 7.18C shows the carrier wave modulated by the 1000 cycle audio tone at a level engineers describe as 100% modulation. That is, the power of the wave is being varied from zero to twice



A. A 4.0 MHz radio frequency wave is displayed on an oscilloscope. Although radio frequency, it has the same waveform as 60 Hz AC, or audio frequencies (50 to 20,000 Hz). There is no modulation.



B. The same 4.0 MHz radio frequency carrier wave is modulated partially by a 1000 Hz audio wave. Note how the power of the wave, or its amplitude, varies at an audio frequency rate.



C. The same 4.0 MHz carrier wave is amplitude modulated at near the maximum level (100%). The result is maximum possible volume in the radio receiver.

Figure 7.18 Radio Frequency Carrier Waves

its unmodulated value. Exceeding this level would interrupt the continuity of the carrier wave and cause the voice or music to be badly distorted.

Review of Amplitude Modulation Principles

In review, the transmissions of a standard broadcast station in the frequency range 1600 kilohertz to 540 kilohertz take place on a constant frequency but with the power of the wave transmitted varying at a voice or music rate. A variant of this type of AM modulation is the *single-sideband* technique: suppressing the carrier and one **sideband** (that part of the wave above or below the horizontal axis of the oscilloscope pictures) and transmitting only a single sideband. This technique is widely used by the military, by commercial communications, by radio amateurs, and recently by citizens band operators. At present it is not suitable for broadcasting music. During reception of a single-sideband signal, a locally generated carrier must be added before demodulation can take place.

Principles and Advantages of Frequency Modulation

FM, or **frequency modulation,** was developed mainly by the late Edwin Armstrong. In the case of FM broadcasting, the carrier wave itself moves alternately to a higher or a lower frequency at a voice or music rate, always returning to the assigned frequency of the station when the program material drops in level. With FM, only the frequency varies at an audio frequency rate. The power of the carrier wave remains constant at all times.

FM has two major advantages. First, most sources of radio interference, such as lightning flashes, automotive ignition, and leaking power line insulators, produce radio waves having an AM characteristic. A receiver designed to respond only to frequency modulation rejects almost all of the interference from these sources. Also, the system of FM transmission and reception lends itself well to the simultaneous generation of another carrier close to the frequency of the main carrier. It may be used to carry the additional program information required in stereo transmission and reception. It is also possible to have more than one of these subcarriers. They can be used for such diverse programming applications as background music for dental offices and instructional programs for the blind.

AM and FM in Television

Television incorporates the technologies of both AM and FM modulation. In the U.S. system the picture information and the controlling pulses that keep the picture tube synchronized with the camera in the TV studio are transmitted by a separate AM transmitter of high power. The sound is transmitted by a separate FM transmitter of considerably less power and operating on a slightly different frequency. The output of the two transmitters is combined and fed into a common transmitting antenna. A TV receiver is partially an AM receiver and partially an FM receiver. They share a number of common circuits, including the power supply.

Only a few other nations use the U.S. system. At Britain's BBC and at television stations in much of Western Europe, both the picture and sound channels are AM in design. Yet, because of well-designed audio systems and husky loudspeakers, the audio quality is excellent. Chapter 10 contains more information about the transmission and reception of AM, FM, and TV signals.

SUMMARY

A transducer is basically a device that changes one type of energy to another. Some good examples are microphones, loudspeakers, and photoelectric cells. Their effectiveness is always less than 100%, and often some of the energy sought is turned partially into heat. Some of the most recently developed transducers are LEDs (light-emitting diodes) and LCDs (liquid crystal displays). They have already found wide application in such devices as pocket calcu-

lators, digital watches and clocks, and digital counters for precise frequency measurement.

Active devices are electronic components that have the capability of controlling electron flow through their structure over a very wide range as a small signal, usually a small voltage change, is applied to their control terminal. The first truly active device was De Forest's triode (three-element tube). It was capable of both signal amplification (the generation of a large signal from a small one) and oscillation (the generation of either an audio or radio frequency signal). The electrons that power the triode tube are liberated from a heated cathode, or filament, by a process known as thermionic emission.

The other two active devices in wide use are the bipolar transistor and the field-effect transistor. Their operation closely resembles that of the triode tube with one exception: Their supply of electrons is derived from their very elements. Both types use silicon doped with either boron or arsenic. This treatment makes free electrons available from the crystalline structure of silicon itself.

Both vacuum tubes and transistors should be considered sophisticated variable resistances. As a varying signal is applied to their control elements, their internal resistance varies widely and permits few or many electrons to flow through their structures and into the external circuits. In effect, the signal applied to their control elements actually controls the large amount of energy supplied by the power supply in the output circuits. The control device requires no power, or very little power, to actuate it, so the weakest signal is sufficient.

For billions of years the earth has experienced radio waves of widely differing intensity and wavelengths caused by natural phenomena such as lightning flashes and static discharges. Only as recently as the mid-1800s was the existence of radio waves predicted, and it wasn't until the last decade of that century that Heinrich Hertz published his discovery of this form of electromagnetic radiation.

At first, experimenters generated radio waves with an electric arc, or spark, as the generating device and used a receiver that employed either metal filings or a crystalline substance (such as silicon) as a detector. Then the Alexanderson alternator was placed in service to generate radio waves of long wavelengths and low frequencies. It made possible for the first time the quality transmission of voice and music. Within a few years De Forest's triode had revolutionized the communications field. It was used both to generate radio waves and to detect and amplify the information superimposed on them at the broadcast station.

Radio waves, like all forms of electromagnetic radiation, move through space at the speed of light—approximately 186,300 miles per second. Surrounding the earth's surface are layers of ionized gases, principally nitrogen and oxygen, which either refract (bend) radio waves or reflect them back to the earth's surface. How much bending or reflection that takes place is dependent upon a number of complex factors: the frequency of the radio wave, the angle of radiation from the antenna, the time of day or night, and the timing of the eleven-year sunspot cycle. Signals that are not reflected or refracted sufficiently by the ionosphere to return to the earth's surface pass through the layers of ionized gas into outer space, where, theoretically, they will continue to move at the speed of light forever.

KEY WORDS AND CONCEPTS

active device
amplification
amplitude modulation (AM)
analog
base
bipolar transistor
detector
field-effect transistor
frequency modulation (FM)
gate
ground wave
ionosphere
modulation
multihop transmission
oscillation
sideband
skip distance
sky wave
sunspot
thermionic emission
transducer
triode

SELF-TEST QUESTIONS

1. In your own words, what does a transducer do?
2. How would you describe a piezoelectric device?
3. What newly developed types of transducers are used in the displays of digital watches and pocket calculators?
4. What two unique functions can all active devices perform?
5. Who is generally credited with the invention of the first three-element vacuum tube?
6. In recent years, electronics engineers have found it convenient to compare the operation of active devices to what common electronic component?
7. Which one of the 92 elements is most widely used today in the manufacture of transistors and integrated circuits?
8. What is the role of an active device in maintaining oscillation of an oscillator circuit?
9. In a conventional oscillator circuit, how can the frequency of oscillation be readily changed?
10. Who is the pioneer in radio and electronics associated with the transmission of the first wireless signal across the Atlantic?
11. What active device replaced Alexanderson's mechanical alternator as a generator of radio frequency waves?
12. In radio frequency communication, what is meant by a sky wave signal?
13. How does the eleven-year sunspot cycle affect transmission and reception in the standard AM broadcast band (1600 kilohertz to 540 kilohertz)?
14. Why were the earliest radio transmissions sent in telegraphic code?
15. What development by a General Electric research team made possible good-quality transmissions of both voice and music by radio communication?

8 Electronics in Medicine and Health Occupations

This pollen grain has been magnified about 3000 times. (Photo courtesy of RCA)

Historical Perspective
- The Curie Family
- Wilhelm Roentgen
- Arsene D'Arsonval

Overview of Medical Electronics
- Operating Room Equipment
- Emergency Room Equipment
- Intensive Care Equipment
- Coronary Care Equipment
- General Patient Care Equipment

Electronic Equipment for Diagnosis
- X-Ray Machines
- CAT Scanners
- NMR Scanners
- Other Scanning Machines
- Ultrasound Equipment
- Electroencephalograph
- Electromyograph
- Electrocardiograph
- Thermographic Equipment
- Electron Microscope

Electronic Equipment for Therapy
- Diathermy
- Ultrasound
- Acupuncture

Electronic Life-Support Equipment

Maintaining Medical-Electronic Equipment

Summary

Key Words and Concepts

Self-Test Questions

This chapter is the first of a series of chapters that describes how electronics is being currently applied. In this chapter, the focus is on medicine and related health occupations. It is directed to workers in patient care, public health, and health research who assist independent practitioners in providing health service.

The first section comprises a brief historical perspective on the contributions certain outstanding researchers have made to our knowledge of bioelectronics and its application to medicine and healthcare. Next is an overview of the use of electronics in a modern hospital or healthcare unit, followed by a detailed account of specific equipment used in diagnostic techniques. Finally, we examine how electronic equipment is used in therapy and describe troubleshooting techniques for medical-electronic equipment.

HISTORICAL PERSPECTIVE

The Curie Family

The first electronics-related breakthrough in medicine occurred in Paris in 1880 when the Curie brothers, Pierre and Jacques, discovered the phenomenon of piezoelectricity in certain crystals, including those of quartz (silicon dioxide) and Rochelle salts (potassium acid tartrate). This same phenomenon is the basis for today's wide use of ultrasonics in diagnostic scanning and therapy.

The Curies found that when certain crystals are stressed mechanically, electric charges appear on their surfaces. Conversely, when stressed electrically by the application of an electric charge to their surfaces, the same crystals move mechanically. This mechanical movement produces waveforms within the range of human hearing (sound waves) and waveforms so high in frequency as to be beyond audibility (ultrasonic waves).

Although this discovery predated that of Roentgen by at least a decade, there was no immediate application for it, so Pierre Curie directed his research first into the field of magnetism and then, in collaboration with his wife, Marie, into a long-term study of radioactivity. Eventually the husband-wife team was joined by their daughter, Irene. Much of the use of the metal radium for radiation treatment in early medical centers was based on the findings of this three-member research team. Both Marie and Irene Curie later died of leukemia, very possibly linked to excessive exposure to radiation from radium and other radioactive sources.

Wilhelm Roentgen

Wilhelm Roentgen (pronounced "Runt-gen"), a researcher in physics at the University of Wurzberg, Germany, was fascinated by the brilliant discharge given off by a Crookes tube when a high voltage was applied across the tube's two terminals. The Crookes tube, or **cathode-ray tube,** was a cylinder of glass with an electrode sealed into place in each end. The tube was pumped to a high vacuum and then sealed off (Figure 8.1). The moment a high voltage was applied, there was a brilliant discharge at the cathode (negative electrode) that the developer of the device, Sir William Crookes, had referred to as a cathode ray.

Not only did the cathode rays produce visible light, they also caused many minerals to fluoresce, or glow. They also penetrated cardboard and thin sheets of wood with ease and even fogged photographic film that had been screened from the visible rays by layers of cardboard. Since Roentgen was unsure of the nature of the radiation causing this

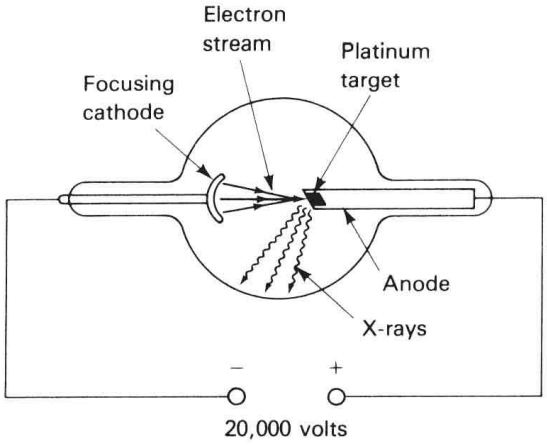


Figure 8.1 Early X-Ray Tube Basically a cathode-ray tube with a platinum target, this early X-ray tube produced a beam with little penetrating power.

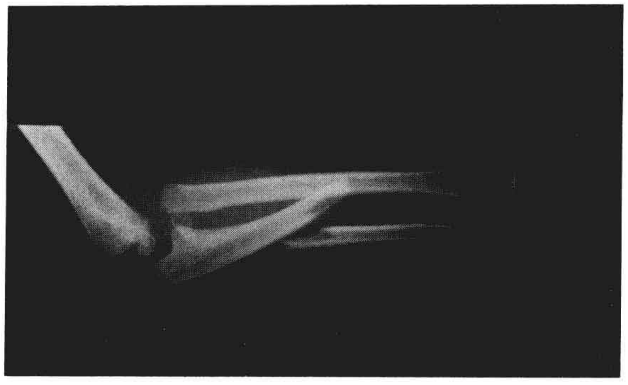


Figure 8.2 X-Ray of Fractured Forearm Very soon after their discovery in 1895, X-rays were employed in hospitals as a diagnostic tool. (Photo courtesy of Ruth McCarroll)

phenomenon, he chose to refer to the rays as **X-rays** (X being a mathematical symbol for the unknown). However, contemporary scientists were beginning to understand that X-rays were just another form of electromagnetic radiation from the vast spectrum that included radio waves, heat waves, infrared and ultraviolet rays (as well as those of visible light), and gamma rays.

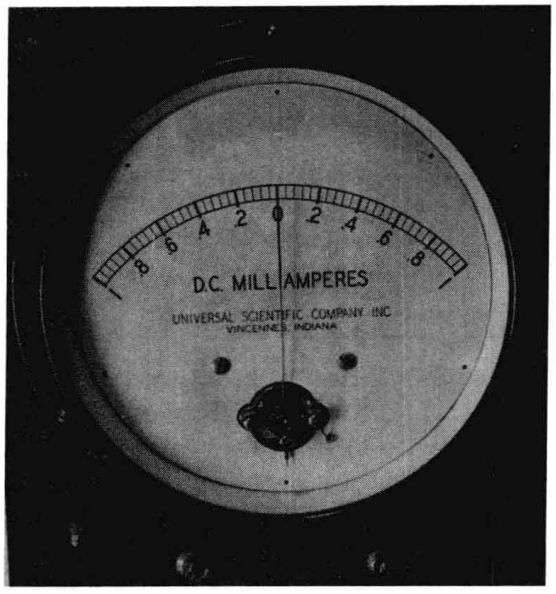
Roentgen became particularly interested in the ability of X-rays to penetrate apparently opaque materials with ease and expose a photographic negative placed beneath them. A year after this discovery in 1895, he displayed in a public meeting an X-ray of the hand of a volunteer from the audience. That X-ray closely resembled modern X-rays. The bones of the hand resisted X-ray penetration and appeared white in the negative, while the tissue of the hand, allowing ready penetration of the rays, appeared gray or near-black.

Roentgen's findings were almost immediately employed in hospitals for locating shrapnel and bullet fragments in the limbs of soldiers brought from the battlefield and for setting of fractured bones (Figure 8.2). In addition to diagnostic applications, X-rays were soon employed in the treatment of malignant tumors and deep-seated infections. Exposure to X-rays was not carefully controlled at first since the dangers of excessive radiation were not fully realized.

Arsene D'Arsonval

Arsene D'Arsonval, a French researcher whose major field was **biophysics** (the application of physics to biological problems), was perhaps the most important figure in the introduction of electronics to medicine and healthcare. D'Arsonval developed the electric meter movement based on a moving coil rotating in the field of a horseshoe magnet. All modern X-ray machines are equipped with a D'Arsonval milliammeter to read the anode current and a D'Arsonval voltmeter to read the anode voltage.

Except for oscilloscopes and digital readout instruments, the D'Arsonval meter is still the most widely used indicating instrument in hospitals and healthcare units. The meter can take a patient's temperature much more rapidly and accurately than can a glass clinical thermometer. The meter can also determine precisely the acidity or alkalinity of a solution. It can match the color of a blood sample against a color standard. It can sample and meter the radiation from an X-ray tube by teaming the



This galvanometer employs the widely used D'Arsonval movement with its center-zero scale. By substituting a tracing pen for the meter's scale pointer, the galvanometer can be used in an electrocardiograph to provide a tracing of heart action.

meter movement with an appropriate transducer to convert what is to be measured into a changing DC current. Once the current has been converted, an appropriately calibrated dial scale allows direct reading in the required units. (An active device amplifier is necessary between the sensor or transducer and the meter, printout, or numeric readout device.)

Another important application of the D'Arsonval meter movement is powering the tracing pens used in electrocardiographs (for heart disease diagnosis) and electroencephalographs (for brain damage diagnosis). In both machines a strip of continuously moving graph paper travels from a supply reel to a take-up reel at a predetermined rate. The amplified electrical impulses from the heart or the brain are fed to the moving coil of one or more D'Arsonval movements (center-scale zero). Attached to the moving coil is a stylus that swings back and forth over the moving graph paper. In some machines the stylus is electrically heated, and chemically treated graph paper must be used to produce a trace. In other designs, the moving arm terminates in an inked pen, and no chemical treatment of the graph paper is required.

D'Arsonval was one of the first researchers in biophysics to point out the electrical nature of neural impulses. He also determined that the difference in potential necessary to set these impulses in motion arose from the nature of living cells and their electrochemistry. D'Arsonval also realized the importance of amplifying electrical impulses so that they could be seen on an electric meter. Since vacuum tubes and transistors had not yet been invented, he used Alexander Graham Bell's new invention, the telephone, to magnify these nerve impulses and was quite successful.

Development of Diathermy

D'Arsonval noted in his experiments with laboratory animals that electromagnetic waves of comparatively high frequency and short wavelength did not produce the characteristic symptoms of electrical shock. Radio waves instead produced a gentle, deep, warming effect on body tissue and were without undesirable side effects. Only when an excessive amount of energy was used did tissue damage result, and it produced burns, not shock. In 1895 D'Arsonval employed his new bioelectronic technique in a Paris hospital to treat muscle aches, arthritis, and circulatory problems. This new form of therapy was first known as D'Arsonvalization and was a well-accepted medical technique by 1910, when it had become known as **diathermy.**

Development of Treatment for Electrical Shock

D'Arsonval's final great contribution to medical technology was his recognition that in many cases victims of electrical shock could be saved if resuscitation efforts were started immediately. The techniques he suggested are still followed today, and probably thousands of electrical shock victims have

been revived because D'Arsonval's recommendations were followed.

OVERVIEW OF MEDICAL ELECTRONICS

Within less than forty years, electronic equipment has had a remarkable impact on treatment and diagnosis in hospitals and healthcare centers. With the exception of X-ray units, the most daring piece of electronic equipment to be found in a typical hospital at the time of World War II was likely to be a diathermy machine. Only large medical centers boasted an electrocardiograph, and usually the entire staff had only one machine available. Today, electronic machines are used throughout hospitals. This section presents a description of key medical-electronic items and their applications.

Operating Room Equipment

In operating rooms, surgeons, anesthesiologists, and assisting personnel keep close watch on specialized **electrocardiographs** (EKGs) and **electroencephalographs** (EEGs) used, respectively, to monitor heart action and brain activity. While conventional EKG and EEG units employ a recording stylus moving over a continuously traveling strip of graph paper, the units used in surgery frequently use the moving trace of an oscilloscope for their display (Figure 8.3). The oscilloscope gives much higher visibility compared to the moving graph of the conventional EKG and EEG. This equipment is frequently mounted in a castered frame in constant view of people working around the operating table (Figure 8.4).

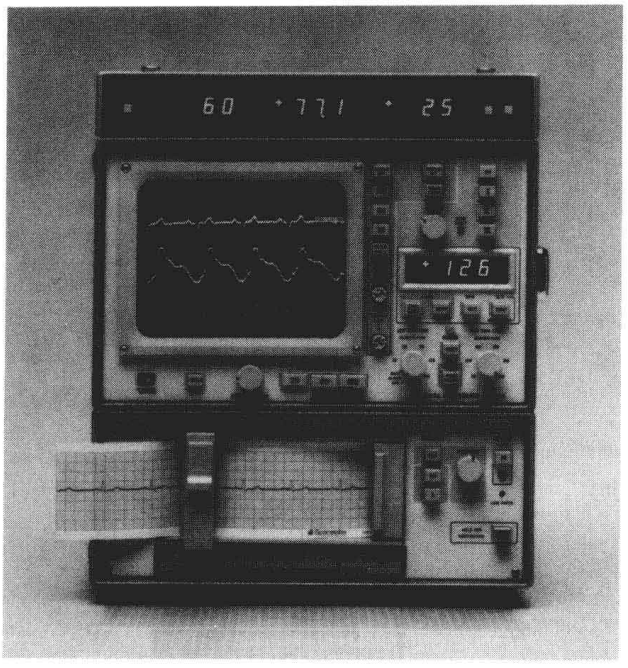


Figure 8.3 Operating Room Monitoring Unit Surgeons and anesthesiologists can quickly determine life signs by a glance at this instrument, which provides digital readout, brilliant scope traces, and one selected printed record. Heart rate, temperature, blood pressure, and EKG tracings are constantly monitored. (Photo courtesy of Spacelabs, Inc.)

Transducers to Measure Blood Pressure

Although not a universal practice, blood pressure may be measured by the use of a small catheter (sterile plastic tube) inserted into a vein or an artery and then brought out to an externally mounted electronic transducer. The transducer transforms the pressure exerted into a varying voltage that can be used to operate an oscilloscope or a digital readout using light-emitting diodes (LEDs). Such blood pressure meters usually work well with normal patients but are suspect if the patient goes into operative shock.

Laser Beams in Surgery

Ophthalmologists use a comparatively new electronic tool, a **laser** (acronym for light amplification by stimulated emission of radiation). Lasers generate *coherent light* (light of a single frequency or wave-

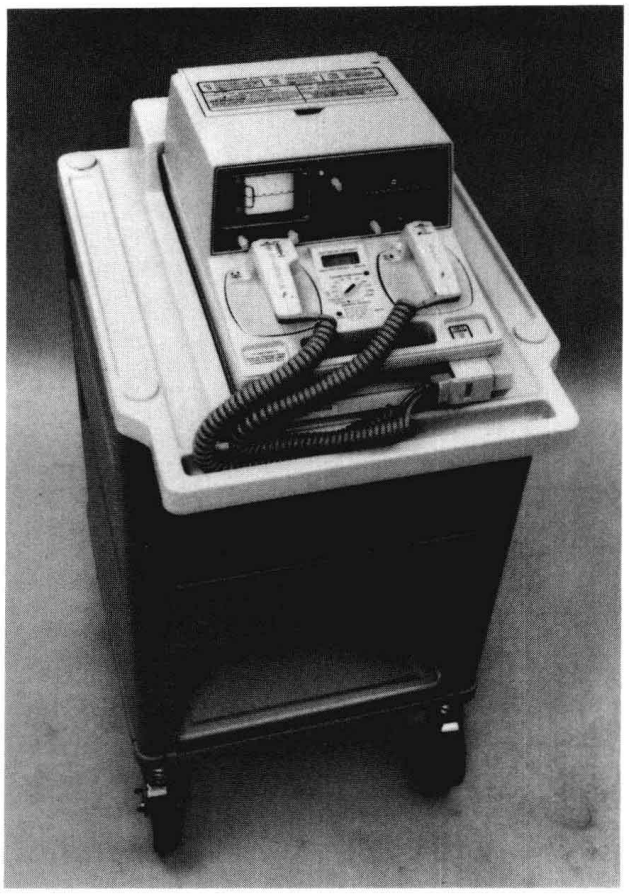


Figure 8.4 Defibrillator Monitor on Accessory Cart This completely self-contained resuscitation system is powered by a storage battery. Drawers of cart hold related equipment and supplies. (Photo courtesy of Hewlett-Packard)

length) by electronic means. Because the light generated is of a single wavelength and in phase, it can be concentrated into a needle-sharp beam of great energy. The ophthalmologist, using the beam in short bursts directed through the lens of the human eye, can electronically stitch a detached retina before it degenerates and causes permanent loss of vision.

Life-Support Equipment

Larger surgeries have heart-lung machines that can support the patient's need for arterial pressure and oxygenation of the blood in certain kinds of surgery, particularly open-heart surgery. Heart-lung machines contain considerable electronic instrumentation. They usually bypass blood and oxygenate it according to the plan for the particular patient. Since such equipment is usually too cumbersome and requires too much time to set up, a portable resuscitation center is used in emergencies. This unit is usually battery powered and contains an EKG preamplifier, cable-mounted electrodes that can be used to start a heart whose action has stopped during surgery, and some form of timing pulse generator. The generator can also be used as a form of heart pacer and is of particular value when heart fibrillation—chaotic electrical activity throughout the heart and loss of regular rhythm—occurs. The pacer overrides random electrical impulses and restores the heart to its normal pulse rate of perhaps 70 beats per minute.

Emergency Room Equipment

The electronic equipment of the typical emergency room closely resembles that of the hospital's operating rooms and surgeries. In emergency rooms much of the equipment is mounted on wheeled racks or carts so that it can be quickly moved from one area to another.

Mobile Units

An extension of emergency room service is offered by mobile emergency care units. In the larger cities ambulances with equipment and personnel capable of professional emergency care have been in operation for many years. The basic equipment usually consists of a resuscitating unit for cardiac and electrical shock cases and a defibrillator with an electrocardiograph whose output is normally displayed on an oscilloscope (Figure 8.5).

From the patient's viewpoint, the most important piece of electronic equipment aboard is probably a two-way radio communications unit that keeps the ambulance crew in constant contact with medical personnel at the emergency room during the drive to the hospital. Radio communication is particularly

Figure 8.5 Detailed View of a Battery-Powered Defibrillator Monitor The two removable electrodes are mounted on panel brackets. An EKG graph is at the left. An EKG waveform display and a digital readout of pulse rate is on the oscilloscope at the right. The unit is readily adapted to use in ambulances. (Photo courtesy of Hewlett-Packard)

helpful when the emergency involves cardiac arrest, where constant observation and treatment en route to the hospital may result in saving a life.

Aircraft as Ambulances

Recently a number of helicopters and light aircraft have been equipped with medical-electronic equipment and two-way radio communications units to perform emergency service similar to that of an ambulance. A helicopter is particularly advantageous because it can quickly transport an accident victim from a remote rural area or the site of a highway crash to the hospital's heliport.

Intensive Care Equipment

Usually located adjacent to the hospital's principal surgeries, the intensive care facility provides continuous observation of patients while they recover from the effects of surgery and anesthetics or from other problems not directly related to surgery. Electronic equipment in intensive care units monitors heart rate as displayed on a special electrocardiograph, blood pressure, and respiration rate. All three readings will probably be displayed on separate oscilloscope traces because of the high visibility of such displays. The monitors for blood pressure and respiration rate require sensors that are basically transducers. They translate the pressure from a vein or an artery (as exerted through an inserted catheter) or from the expansion and contraction of the patient's chest wall into an electrical signal that produces a typical display on an oscilloscope.

Usually the intensive care unit is organized so that electronic signals from each patient's EKG preamplifier and various bodily function sensors can be brought through conduit to a master monitor unit located at a central desk. Here a highly skilled nurse can keep careful watch over the displays on the various oscilloscopes, interpret these patterns, and call for immediate assistance if an emergency arises.

Special multitrace oscilloscopes are used so that several types of information can be displayed simultaneously on a single, compact monitoring unit that simplifies the task of the person assigned to study the constantly changing traces. In addition to the specialized scopes is some form of electronic alarm system that sounds a buzzer or rings a bell at any sudden change in a patient's monitored signals. In medical-surgical intensive care units, more emphasis is being placed on providing patient care at bedside. Under these conditions, central monitoring performs an adjunct, or backup, service.

Coronary Care Equipment

In larger hospitals where considerable specialization is practical, coronary (heart) patients are usually treated in a specially equipped and staffed area called a CCU (coronary care unit). The electronic instruments used are particularly appropriate to cardiac monitoring—for example, an electrocardiograph with an oscilloscope across its output and additional digital readout displays that provide numerical monitoring of heart rate and blood pressure. (These displays must be actuated by suitable sensors, or transducers.) The

staff also uses a computer to detect subtle changes in an EKG that would indicate *arrhythmias* (variations from normal heart rhythm). In many CCUs, much of the monitoring equipment is centrally located so that a specially trained nurse can watch for life-threatening arrhythmias. Multitrace oscilloscopes are widely used to observe all the patient's life signs on a single instrument.

Cardiac patients who are no longer in immediate danger of another heart attack must still be constantly monitored for physiological changes that might indicate a possible recurrence. Therefore, these ambulatory patients are permanently wired to an ultracompact electrocardiograph whose output feeds a limited-range FM radio transmitter. Both units are powered either by disposable dry batteries or by rechargeable units of the ni-cad type. A multiple-channel FM receiver at the central monitoring position picks up the individual transmissions, separates the patient's health information from the carrier wave, and processes the data, which are then displayed on the screen of an oscilloscope. Additional electronic instrumention using an LED diode display provides a visual presentation of the patient's heart rate.

This type of monitoring is called *telemetering*, which means to conduct measurements at a distance. The nurse in charge may employ manual switching to select the individual to be checked, or the equipment can electronically scan the individual signals in automatic sequence. An electronic warning system is usually tied into the circuits at the central monitoring position. It will sound a buzzer or ring a bell at any sudden change in any patient's monitored signals beyond predetermined limits. Also at the central desk is the versatile multitrace oscilloscope with one scope assigned to each patient's signals.

General Patient Care Equipment

Electronics has made hospital stays considerably more pleasant—for example, by making available cable television and music. Electronics has also improved items connected with general patient care. The glass thermometer, which required several minutes to produce an accurate reading, is being rapidly displaced by an instrument that uses a **thermistor** as a sensor and gives an instant and highly accurate reading (Figure 8.6). A thermistor is an electronic device whose resistance varies with temperature changes.

When a thermistor is connected in series with a small battery and a sensitive meter with a properly calibrated scale, all the basic circuitry of an instant-reading clinical thermometer is available. The thermistor itself has to be enclosed in epoxy or its equivalent since it must be impervious to the moisture of the patient's mouth and to the alcohol used for sterilization.

Electronics has also greatly reduced the time consumed by preoperative blood tests. Once tedious and time-consuming, these tests can now be done by a minicomputer that watches over signals flowing from electronic sensors in contact with tiny amounts of blood and blood serum. In just a few minutes all the tests are completed, and the computer prints out the data on a special form (Figure 8.7). The patient saves in two ways: The tests are performed quickly, and costs are greatly reduced. In some cases a patient can enter the hospital for minor surgery and go home the same day.

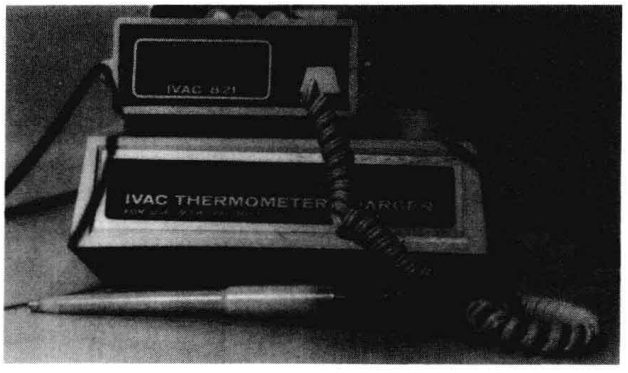


Figure 8.6 Electronic Clinical Thermometer The thermistor action provides a highly accurate reading in seconds. (Photo courtesy of Sacred Heart General Hospital, Eugene, Oregon)

Figure 8.7 Autoclinical Analyzer for Complex Tests These units have greatly reduced the time consumed by preoperative tests. (Photo courtesy of Sacred Heart General Hospital, Eugene, Oregon)

ELECTRONIC EQUIPMENT FOR DIAGNOSIS

Diagnosis means the gathering of as much medical information about the patient as the attending physician needs to make informed decisions about the patient's medical problem and its treatment. Diagnosis begins with recording the patient's temperature, weight, and blood pressure; listening to heart and lungs with a stethoscope; and visually examining the retina of the eye. The physician will probably also order blood tests and urinalysis as a routine procedure. After the routine tests have been performed, electronic diagnostic equipment provides additional, highly specialized information in the form of X-rays or electrocardiograph readings.

X-Ray Machines

Modern X-ray machines are much the same in design as that developed by Roentgen in the 1890s. A glass cylinder is flared in the middle and houses the cathode and the anode. The glass contains enough dissolved lead dioxide to block X-ray radiation in all directions except downward. A window of special glass is set into the lower surface of the bulb to permit the ready passage of X-rays.

As shown in Figure 8.8, a filament of tungsten wire is recessed into one electrode, known as the cathode. When this filament is heated so it glows, an abundance of free electrons is liberated from its surface. At the opposite end of the tube is a large copper electrode, known as the anode, which serves as the target for the free electrons liberated from the cathode. The end of the copper anode facing the cathode is angled downward so that the X-rays generated will be directed down through the window of special glass and beyond the tube. Usually a small rectangle of tungsten is set into the angled surface of the copper and acts as a focal point for the beam of electrons arriving from the cathode.

Before the tube can be placed in operation for the first time, it must be pumped to a very high vacuum and sealed off with a gas jet. The cathode operates substantially at ground potential (essentially no potential or voltage), while a very high DC voltage is applied to the anode—perhaps 150,000 volts on the average. Free electrons, accelerated to a speed that would whisk them from New York to Chicago in a second or less, strike the anode with a terrific impact and momentarily force the electrons of the tungsten atoms out of their normal orbits. When the electrons snap back into position, X-rays are generated. X-rays are just one of the many forms of electromagnetic waves that include radio waves, heat waves, light waves, and gamma and cosmic rays. X-rays are shorter in wavelength than ultraviolet light, longer in wavelength than gamma rays.

To a limited extent, the same X-ray machine can be adjusted to produce both "soft" and "hard" X-rays. Longer-wavelength (softer) X-rays are about the same frequency as ultraviolet light and have limited penetration power. Shorter-wavelength (harder) X-rays approximate the frequency of gamma rays and penetrate the human body with relative ease. Raising the anode voltage and the electron beam current tends to shift the X-rays emitted by the tube from softer to harder.

Unlike light rays, X-rays are not readily focused by a lens system. Instead, they spread out from the

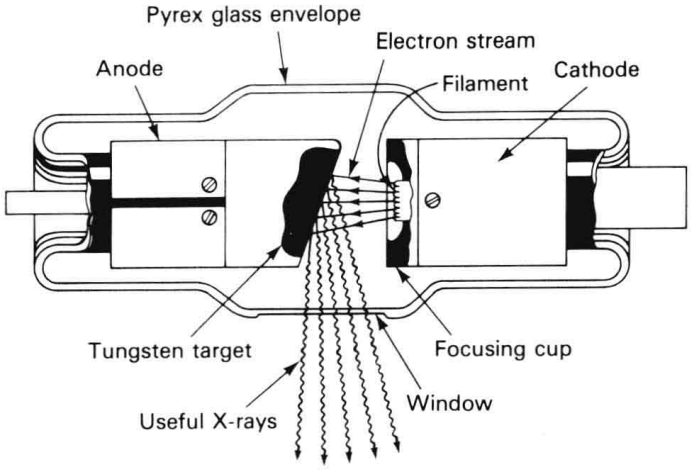
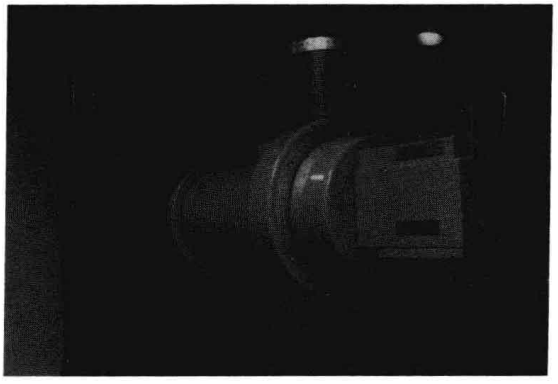


Figure 8.8 Modern X-Ray Tube Typical of the X-ray tubes manufactured today, this unit produces a powerful, sharply focused beam of great penetrating power. (The necessary cooling system is not shown.)

focal point on the target of the anode to form an elongated, narrow cone resembling a flashlight beam. Therefore, the tube must be quite close to the patient's body for the photographic negative to be sharp and clearly outlined.



This conventionally designed X-ray machine of the orthovoltage type can be used for applications other than radiation. (Photo courtesy of Sacred Heart General Hospital, Eugene, Oregon)

Normally the patient lies on a wooden or plastic table under which a large section of special negative film is mounted in a holder. A common size for a chest X-ray, for example, would be about 14 inches by 17 inches. As the X-rays enter the body, they are almost stopped by bony structures that appear white in the film negative. Normal tissue, the various organs, deposits of fat, and air bubbles all absorb X-rays to varying degrees and, in turn, expose the negative film below the patient to variable amounts of radiation. Ordinarily, organs such as kidneys, gall bladder, and liver are not recorded clearly unless the patient has been previously injected with a dye that concentrates in these areas. The use of the so-called barium milkshake to outline the upper gastrointestinal tract is a common practice.

Control Equipment

In all modern installations, the X-ray machine's control equipment has been moved well away from the instrument itself because of the danger to the operator of an accumulated dosage of radiation over a long period of time. Radiation accumulation has definite carcinogenic (cancer-producing) possibili-

ties. The control console always contains D'Arsonval meters—one of about 0 to 150,000 volts to measure the anode voltage, another of about 0 to 50 milliamperes to measure the electron beam current. Also included is a very precise electronic timer to ensure an exact amount of radiation. Fail-safe electronic circuits are also included and will shut down the instrument if the timer fails. Otherwise, dangerous overradiation of the patient might result.

Because the efficiency of producing X-rays by impacting an anode with a high-velocity beam of electrons is so low, about 99% of the electron beam's energy is turned into heat energy. Only about 1% becomes X-rays. Particularly with large tubes, elaborate ventilation systems and heat exchangers are necessary to get rid of the unwanted heat, which might cause damage to the equipment and discomfort to the patient. Both forced air and circulating oil systems are commonly used for cooling.

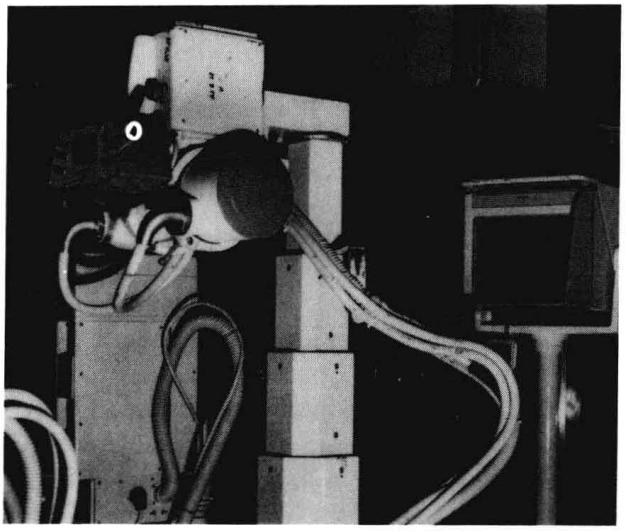


Figure 8.9 Modern Fluoroscopic X-Ray Installation A TV monitor for imaging the instrument's output isolates the viewer from possible X-ray radiation. (Photo courtesy of Sacred Heart General Hospital, Eugene, Oregon)

Fluoroscopy

In addition to his many other inventions, Thomas Edison should be credited with inventing the technique of **fluoroscopy.** Taking advantage of the fact that certain chemicals fluoresce, or glow, when struck by X-rays, he coated a translucent screen with these substances. Placing the screen behind an individual facing a source of X-rays, he could readily see the action of the person's heart and lungs, the movement of limbs, fingers, and toes, and under certain circumstances the churning of the stomach and its contents. In effect, he had produced a crude motion picture that was to provide helpful insights into the functioning of the human body and its organs.

Because he was unaware of the seriousness of overradiation, Edison was seriously endangering the lives of both patient and professional observer. Today this danger is minimized by using an X-ray source of low intensity, by limiting the fluoroscopy to short periods, and by incorporating into the special screen materials that effectively prevent X-rays from reaching the observer. Figure 8.9 shows a modern fluoroscopic X-ray machine.

Modern electronics uses light amplifiers to provide an image of such brightness that black-and-white pictures can be taken with an ordinary 35 millimeter camera or with a camera using larger-format film. A television camera can be used to feed a video signal to monitors in nearby classrooms. At present, fluoroscopy in color seems to be a distinct possibility, probably through the use of digital techniques. Many physicians prefer the use of 16 millimeter motion pictures, which allow them to study and restudy functions such as heart and lung action.

Dental X-Rays

Another electronic device used in diagnosis is the dental X-ray machine. Although the principle of its operation is the same as that of the instruments used in purely medical practice, the dental X-ray machine differs in certain particulars. Since it is always used at close range—say, 6 to 8 inches—it need not be too powerful, so most machines are scaled-down versions of the units in hospitals or

clinics. At a distance of 8 inches, the anode voltage used will be about 65 kilovolts (65,000 volts), and electron flow from cathode to anode will approximate 10 milliamperes. The incisors take the least exposure for a satisfactory negative. The more massive molars take appreciably longer. Metal fillings, being completely opaque to X-rays, appear white on the negative, as in Figure 8.10. The teeth, because their enamel coating is quite opaque, appear light gray. The rather porous jawbone passes X-rays rather readily and has a dark appearance with some patterning. A dental assistant can readily identify areas of decay, an abscessed root, or evidence of root-canal surgery.

Because of increased awareness of the dangers of excessive radiation, dentists and technicians no longer hold the X-ray negative material with their fingers as the exposure is made. A bite-wing cardboard holder supports the 30 by 40 millimeter negative firmly in place. Dentists have also reduced radiation by X-raying just a few teeth a year, usually in planned rotation.

Figure 8.10 Dental X-Ray Amalgam fillings show white in contrast to tooth structure and jawbone. (Photo courtesy of Lane Community College, Eugene, Oregon)

CAT Scanners

In the mid-1970s, medical-electronic engineers demonstrated a new X-ray device that united existing X-ray technology to the technologies of the digital computer and of modern television. A breakthrough in the diagnostic field, this collection of X-ray and purely electronic equipment is referred to as a **CAT scanner** (Figure 8.11). CAT is an acronym for computerized axial tomography, the process of making a large number of individual X-rays in rapid sequence. The word *scanner* is best explained by the analogy of television scanning.

It has been recognized for well of over fifty years that in the transmission of a picture over an electronic system, it is impractical to transmit the entire scene in one electronic burst of operation. Instead, the total scene is dissected into bits and pieces by the TV camera and reassembled by the scanning circuits of TV sets. In the TV camera, a

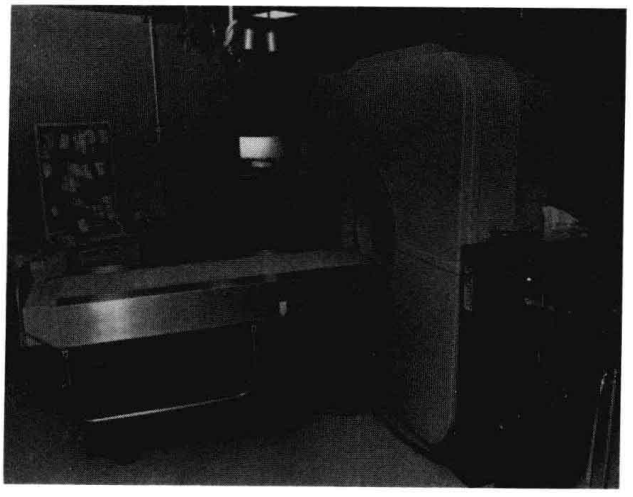


Figure 8.11 Computerized Axial Tomographic Body Scanner (CAT Scan) Multiple X-ray exposures are reconstructed by a computer to provide cross-sectional images of the body. (Photo courtesy of Sacred Heart General Hospital, Eugene, Oregon)

lens brings the studio scene into sharp focus on a rectangle of mica covered with millions of silver globules. Since silver is photoactive, the globules in the light areas of the camera scene tend to lose electrons and become positively charged, while those in dark areas are unaffected.

An electron beam is then made to move systematically across the sensitized area. Starting in the upper left-hand corner of the scene, it moves in a straight line from left to right, discharging the silver globules with its passage and producing a pulsating DC current in the camera tube's output circuit that can be used to reconstruct all the information on that particular scanning line in a television receiver. At the end of the first line, the electron beam is temporarily extinguished while it is moved slightly downward and to the left-hand margin to start the next scanning line. When the bottom of the camera image has been reached, the beam once more is momentarily extinguished and then positioned at the left-hand margin at the top of the camera image, ready to start the entire sequence of scanning lines again.

A TV picture tube with its electron beam in exact synchronism with the camera beam can recreate the camera scene scanning line by scanning line. **Scanning**—the process of acquiring and storing information electronically as the sensing apparatus moves progressively over an area from left to right and from top to bottom—is the basis not only of computerized axial tomography but also of electronic scanners based on radioactivity or on ultrasonic waves.

The CAT scan approach to X-ray diagnosis stems from inherent inadequacies in the conventional X-ray process. Even a highly skilled X-ray reader frequently has difficulty interpreting the two-dimensional shadow pictures taken with an X-ray source above the patient and a sensitized film negative below. Various organs often do not appear clearly, the boundaries between them are hard to visualize, and sizable tumors both benign and malignant may be completely overlooked. These deficiencies occur because in effect, we are looking at a whole sequence of X-ray photographs, each one in reasonably sharp focus but blurred by those in front and those behind it. X-ray tomography resolves this problem.

Tomography

Tomography refers to the X-ray technique of taking a large number of pictures in a scanning pattern and then assembling them in such a manner that structures before and behind the section under study do not show. X-rays cannot be focused with ordinary glass lenses (an optical impossibility), but by the use of a properly programmed digital computer in conjunction with an oscilloscope or a specialized TV monitor, highly detailed pictures of both brain and body can be obtained. The information obtained is often superior to that acquired through exploratory surgery, and it is obtained without the shock and trauma that surgery entails.

The first X-ray tomograph was a comparatively small unit designed specifically for X-rays of the head. Construction of a unit that would perform a full body scan posed new problems, and satisfactory models were not available until the mid-1970s. The tomographic body scanner is probably the largest medical-electronic machine in the hospital, and certainly one of the most expensive. Installed cost may run as high as $1 million.

Construction of a CAT Scanner

A very large C-shaped bracket, pivoted at its center, supports the X-ray source at one end of its length and about 30 sensors made of sodium iodide crystals at its opposite end. The bracket is slowly turned by an electric motor. The cot supporting the patient's body is at the same level as the pivot point on the C bracket so that both the X-ray source and the sensing units around the patient's body can be rotated at a predetermined rate and scanning pattern.

When the X-ray source moves in one direction, the sensors move in the opposite direction. The recorded image is clear because one plane of the picture is kept in focus while all others are blurred by being out of focus. As the X-rays from the source pass through the patient's body, they are absorbed

to varying degrees by bony structures, cartilage, fat, and tissue. Emerging from the patient's back, they are detected by the crystal sensors that serve as transducers to change the X-ray energy into light energy and produce flashes of light (scintillation). Light amplifiers amplify and measure the level of these flashes, and after further electronic processing, the data are fed to the digital computer.

The normal direction of scanning is *transverse*—that is, at right angles to the patient's spinal column. During each scan of the patient's body, a slice, or cross section, of the body's absorption of X-rays from the sharply focused radiation source is sampled, amplified, processed, and recorded in magnetic form by the computer. One scan is completed in about ten seconds, and then another scan can be started. At least eight transverse cross-sectional scans can be completed before the patient must be moved to a new position. Thousands of X-ray images are recorded within a period of just a few minutes, yet the total radiation received by the patient is about the same as would be received in a conventional series of X-ray photographs.

Role of the Computer in CAT Scanning

The digital computer used with the CAT scan is not just a helpful accessory. It is the most important component of the system. As soon as the scan of the desired body area is completed, the X-ray equipment, the sensors, and the signal processors are turned off. Then the computer, directed in its operations by specially written instructions set forth in computer language and known as a *program*, recovers all the information from the scans it has stored magnetically. Storage of a varying magnetic pattern on an iron oxide coating may be either on conventional magnetic tape or on so-called floppy discs, which resemble phonograph records.

The recovered information is then synthesized into the equivalent of a three-dimensional picture that can be used for observation and diagnosis by the attending physician. Many organs and structures that were difficult to see in a conventional X-ray photograph can be seen with remarkable clarity and definition on a scan. The image can be viewed either on a specialized oscilloscope or on a modified TV monitor if a more detailed view is required.

In addition to having one output channel to provide the electrical signal necessary to operate the oscilloscope or TV monitor, the computer always has at least one additional output providing a signal to operate a line printer. Much like an electric typewriter, the line printer records the data from the scans on paper in a form known as *hard copy*. Because both the visual and printed data are recorded so compactly, they can be stored in a small carton or envelope.

Cost of Scanning

The CAT scanner works rapidly and records its data in permanent form that can be readily replayed on other equipment. Therefore, the cost is not out of line with other X-ray diagnostic procedures, particularly in hospitals that tightly schedule the scanner's operation so that it is constantly in use throughout the day.

One of the disadvantages of a CAT scan is that X-rays play an essential role, and X-rays are invasive—that is, penetrating and capable of causing damage to body tissue even when the radiation is carefully controlled. Also, X-rays may be carcinogenic (cancer-causing) in certain cases.

NMR Scanners

Medical science's most recent development is the **nuclear magnetic resonance scanner** (Figure 8.12). In the NMR scanner, X-rays are replaced by an extremely strong magnetic field that surrounds the patient. This magnetic field has no adverse effect on the patient. However, it does line up the magnetic fields within the atoms that make up the body's tissue and bones. (These tiny magnetic fields surround the nucleus of such elements as phosphorus and hydrogen, which have an odd number of protons and neutrons in their central core.) The phenomenon is similar

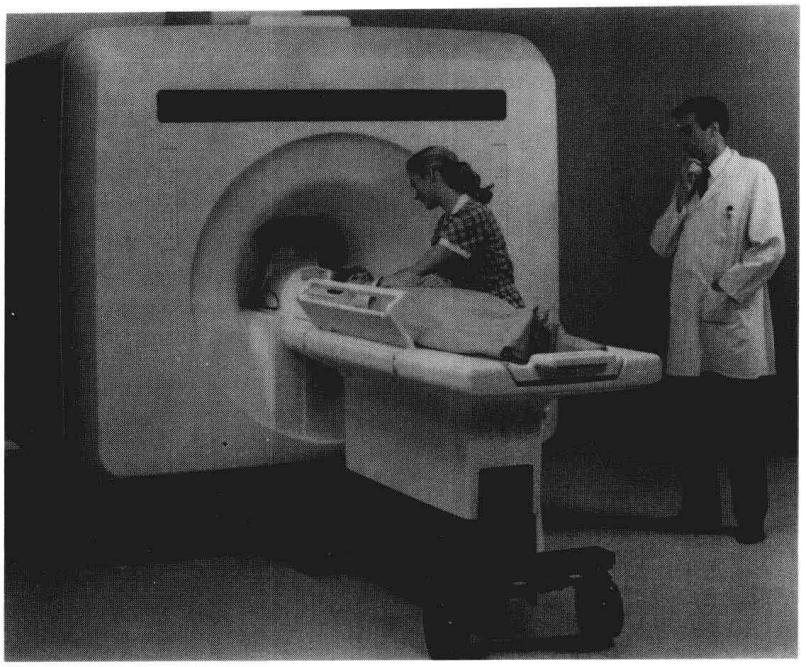


Figure 8.12 NMR Scanner A product of the Medical Systems Group of General Electric, this nuclear resonance scanner uses electromagnetic fields rather than X-rays to scan the patient's head and body. Its computer has disc storage for more than 1000 images. (Photo courtesy of General Electric, Medical Systems Group)

in principle to the lining up of iron filings in the field of a permanent magnet.

In the process of imaging, the NMR scanner also supplies a rapid series of electromagnetic radio frequency pulses that produce a periodic magnetic field at a 90° angle to the primary one. In response to this second magnetic field, the nuclei are turned a quarter-turn from their original direction. They snap back to their first alignment when the pulse terminates. This sudden return to the original orientation produces a magnetic signal characteristic of the particular tissue or bone that contained the atoms and detectable by electromagnetic sensors traveling over a preplanned scanning path. The sensors' output reflects the highly individual nature of the tissue or bone where the changing magnetic signal originated. The output of the sensors is then processed by a computer and restructured into a two-dimensional image in full color on the screen of a TV monitor.

The NMR scanner is a complex and very expensive instrument still in its developmental stage. However, it is far superior to the CAT scan for brain studies since it penetrates the human skull with ease. It is also excellent for examining the chest cavity, blood vessels, and blood circulation and for observing fetal development and growth. The high cost of installing an NMR scanner (at least $1.5 million) will limit its use to the largest healthcare centers.

Other Scanning Machines

Several medical-electronic machines used in hospital and healthcare clinics employ scanning techniques and radiation detection systems closely resembling those of the computerized axial tomograph but are far simpler in design and operation. In the case of the CAT scanner, the radiation source is a sharply focused X-ray beam. In the simpler body-scan machines, the radiation source is radioactive materials introduced into the patient's body. This technology is actually a form of nuclear medicine.

Basics of Radioactivity

Although a detailed study of radioactivity and nuclear medicine is well beyond the scope of this book, some basic principles can be readily understood. Radioactive materials in their natural form (e.g., uranium) or in artificially created forms (e.g., an isotope, or nuclide) spontaneously disintegrate into simpler atoms. In this process of fission, three distinct forms of radiation are given off. The physicist Ernest Rutherford named them *alpha, beta,* and *gamma*. The first two are best described as particles, while gamma rays are a purely electromagnetic form of radiation that have neither mass nor charge. Beta particles are electrons. An alpha particle is equivalent to the nucleus of the helium atom with its two protons and two neutrons. An alpha particle has thousands of times the mass of an electron.

Of these three forms of radiation, alpha particles, because of their greater mass, travel the most slowly and achieve the least penetrating power. Beta particles (electrons), because of their far smaller mass, travel much more rapidly and penetrate more deeply than alpha particles. Gamma rays, being a form of electromagnetic radiation and not particles, have by far the greatest penetrating power. They can penetrate up to 30 centimeters of iron. (A centimeter is about 0.4 inch.)

Detecting Radiation

As radioactive substances decay, all three forms of radiation may be given off at the same time. Consequently, the medical-electronic equipment that employs radiation must be designed to respond to one, two, or all three forms. The design usually takes advantage of the fact that certain crystals—usually those of either sodium iodide or zinc sulfide—act as transducers that readily transform the various types of radiation energy into flashes of light. In this function the crystals are known as **scintillation detectors,** and the principle on which they work is **luminescence** (Figure 8.13).

Radioactive Scan Procedures

A thyroid scan provides a simple illustration of the procedures involved in a radioactive scan. First the patient receives an injection of radioactive iodine isotope (the preferred word is now *nuclide*). After an hour, during which the thyroid gland absorbs the iodine from the bloodstream, the patient is placed under the scanning detector with the unit pressed closely against the neck. A motor is started and scanning commences. In about fifteen minutes the scan is completed, the scanning head moved to one side, and the recorded chart examined. Immediately

Figure 8.13 Computerized Body Scanner of the Scintillation Type Output of scanner (over table at right) is fed as an electrical signal to the computer for imaging on monitors to the left. (Photo courtesy of Sacred Heart General Hospital, Eugene, Oregon)

apparent is the butterfly-shaped outline of the thyroid gland. Within this outline are hundreds, perhaps thousands, of individual dots made by an electrically charged stylus contacting chemically treated graph paper. If the chart shows many areas with comparatively few dots where a dense dot pattern would be the norm, the diagnosis is below-normal thyroid activity.

The entire body can be scanned by the same basic method used in the thyroid scan. First, a suitable radionuclide is injected or swallowed. Certain organs absorb this radioisotope more readily than others. The area immediately above the organ in question is then scanned mechanically with a detection device made up of one or more collimators, crystals, preamplifiers, linear amplifiers, and processors. The resulting electronic data can be displayed on a special oscilloscope and stored in magnetic form in a computer for later printout.

Ultrasound Equipment

As previously noted, the Curie brothers discovered well before the turn of the century that certain crystals vibrated mechanically when subjected to electrical stress and generated corresponding electrical voltages on their surfaces when stressed mechanically. Such crystals were called piezoelectric. Their first major application at supersonic or ultrasonic frequencies took place during World War II, when vibratory waves 20 kilohertz or higher were used in a sophisticated submarine detection technology known as *sonar*. Not until nearly 1950 did researchers develop equipment that permitted the use of **ultrasound** as a diagnostic tool.

Sound waves, which are closely related to ultrasonic waves, travel at differing speeds through media of varying densities. They travel through dry air at approximately 1100 feet per second, through water at about 4700 feet per second, and through steel at a rate in excess of 15,000 feet per second. The relationship between the speed of propagation (and reflection) and the density of the medium is the basis of the use of ultrasonics in medical diagnosis.

When a suitable piezoelectric crystal is placed in contact with the human body and excited by a suitable generator of ultrasonic energy, ultrasonic waves penetrate the body at a rate usually assumed to be about 1500 feet per second in soft tissue (roughly a velocity of 500 meters per second). As soon as the penetrating waves encounter tissue of greater or lesser density than that in which they are traveling, a certain percentage of the energy is returned to its source as an echo. Since the piezoelectric crystal used to generate the ultrasonic waves is an excellent transducer, it is capable of turning the mechanical energy of this returned echo into a corresponding electrical wave. This wave, after amplification and processing, can be used to produce a display on the screen of an oscilloscope. Ultrasonic energy is not introduced into the patient's body continuously but rather in the form of very brief bursts, or pulses, that usually occur at the rate of about 1000 per second.

Because this technique involves pulsing, waiting for an echo, and then pulsing again, it is commonly referred to as a **pulse-echo diagnostic procedure.** The crystal unit—usually barium titanate or lead-zirconium titanate—actually performs two functions in its role of transducer. First, when excited by a pulse of energy from the ultrasonic generator unit, the crystal generates the ultrasonic wave, which is directed into the patient's body. Then, during the time between pulses, electronic switching circuitry makes it possible for the same crystal to "listen" for any returning echoes and convert them into an electrical signal. An electronic timing circuit once more supplies a burst of energy from the ultrasonic generator unit.

Using a single crystal for both generation of the pulse and subsequent detection of the echoes simplifies the basic design of the unit and improves its accuracy. With just one crystal unit in use, the location of the wave's transmission into the body tissue and the location of the wave's subsequent reception as an echo are always at precisely the same point.

Holding the crystal transducer immediately over the area being studied (e.g., over the heart) is just one diagnostic technique that can be used with ultrasound. Ultrasound can also be used in a progressive scanning technique in which a comparatively large area of the patient's body is mapped. This type of scan requires mechanical movement of the crystal

transducer in a manner much the same as that used by the CAT scanner. The same ultrasound scanning technique is used to perform a *B-scan,* which provides excellent two-dimensional views of the human brain without any danger to the patient.

When the B-scan ultrasound signal is fed to the terminals of an oscilloscope after processing, there may be a problem in visualizing the pattern seen on the instrument's screen. A better but far more expensive technique is to feed the processed electrical signal into the input of a small computer. The computer provides additional signal processing plus data storage. The output of the computer can then provide a three-dimensional aspect to the signal and greatly improve ease of interpretation. The computer can also recall all scan data in order. The net result is a "picture," displayed either on an oscilloscope or a TV monitor, that shades from white through intermediate gray to black and has a feeling of perspective. It can be constantly replayed in the entirety of the scan.

Although ultrasound has been used in diagnosis since the early 1950s, since 1970 it has become much more sophisticated because of improved electronic technology, particularly use of the computer. Moreover, ultrasound can perform many diagnostic tasks formerly assigned to X-rays with little or no risk to the patient. For example, ultrasound can be used to check fetal development at regular periods when X-rays would pose serious danger to the unborn child. Ultrasound may also provide more detailed data on the existence and extent of tumors than do conventional X-rays. Ultrasound can also be used very effectively and without danger in examination of the eye for detachment of the retina or for blood clots that are the result of hemorrhage. In this application the crystal transducer is usually hand-held in a fixed position, and scanning is not used.

A specialized ultrasound diagnostic technique used to display both heart structure and motion, as well as heart abnormalities, is known as **echocardiography.** It is especially effective when the imaging is assisted by a computer. Echocardiography accounts for nearly one-third of all examinations in which ultrasound is used.

At present, no industrywide standard exists among manufacturers of ultrasound medical equipment for the exact vibratory frequency used in pulse-echo diagnosis, but it usually falls between 2.0 and 15.0 megahertz. With this exception, the basic design of diagnostic ultrasound equipment is fairly uniform.

As new applications for ultrasound are developed and more machines placed in operation, more ultrasound technicians must be trained. As many as 75,000 new ultrasound technicians will be required by the mid-1980s.

Electroencephalograph

The **electroencephalograph** (EEG) is an electronic instrument that measures and records the small, varying electrical impulses generated by the nerve cells in the brain. When two or more electrodes are placed in contact with the surface of the brain, a minute, varying voltage can be measured between them. Until active devices (vacuum tubes and transistors) that would provide a high order of voltage amplification were developed, the nature and extent of these changing potentials could not be accurately measured.

Today, sophisticated solid-state amplifiers, with low hum and noise levels, can readily measure, process, and display these so-called brain waves. Because applying the measuring electrodes directly to the exposed surface of the brain would involve surgery and its attendant risk, a noninvasive approach is used. The sensing electrodes are applied to the scalp. Intervening bony structure and scalp tissue reduce the electrical potentials to perhaps one-third the level of direct measurement, but high-gain amplifiers easily make up for this loss. The actual potential range encountered can be expected to vary from a few microvolts up to a maximum of about 200 microvolts. (A *microvolt* is 1/1,000,000 volt.)

Conventionally, the EEG has an eight-channel amplifier followed by an eight-channel driver amplifier whose function is to move the eight D'Arsonval recording pens across the constantly moving strip of graph paper. A typical EEG recording setup would employ six pairs of scalp electrodes, arranged

in a standardized pattern, plus a reference electrode that is usually attached to one ear.

The eight individual amplifier units are wired as *differential amplifiers*. That is, their inputs are so arranged that they only amplify the difference in potential between the electrodes of each pair. The resulting output waveform is both symmetrical and irregular. The frequency of the waveforms will usually fall between about 4 and 13 hertz. Interpretation of the waves is beyond the scope of this text, but two examples of normal brain waves are included in Figure 8.14.

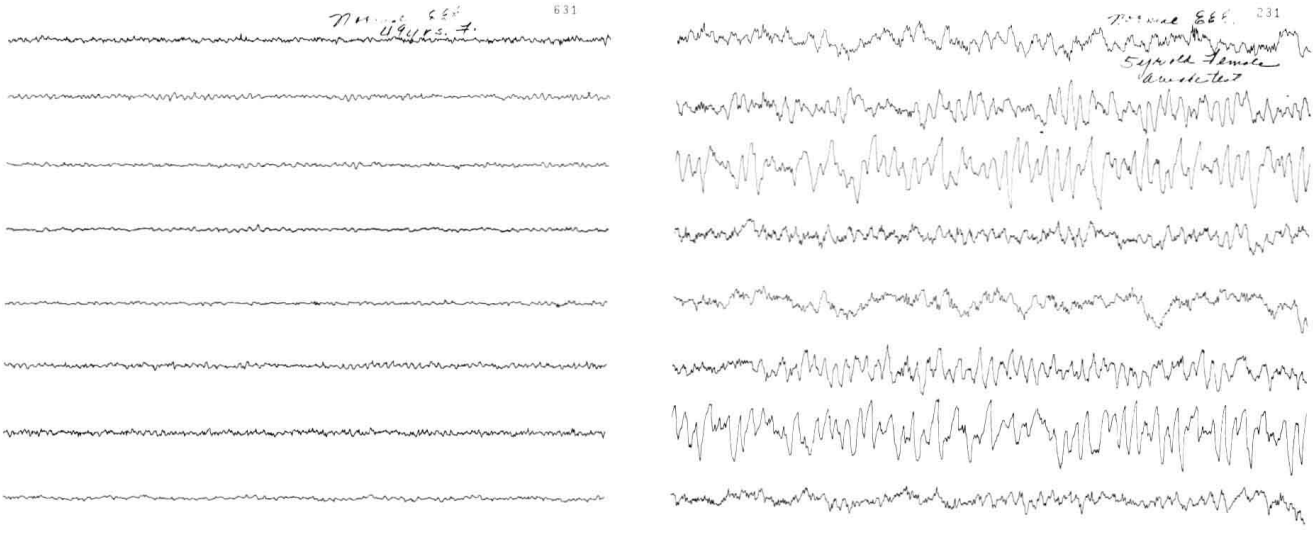
Brain waves are usually produced by tracing pens on a moving sheet of graph paper. Another option is a multiple-trace oscilloscope to present the output of a number of channels on its screen at the same time. Current electronic technology is sufficiently advanced to make it possible to display all eight channels simultaneously if the scope uses special electron-beam sweep circuitry and its screen is coated with a high-persistence phosphor. However, for routine EEG studies, the preference seems to be the traditional presentation on graph paper.

The EEG is a widely used diagnostic instrument. Interpreted by a trained specialist, it can diagnose the extent of brain injury following an accident and determine the probability of a brain tumor or the presence of epilepsy.

Electromyograph

The physical action of muscles contracting is always accompanied by corresponding electrical activity. The **electromyograph** (EMG) picks up, amplifies, processes, and displays this information on the screen of an oscilloscope. An oscilloscope is used for data display because a range of frequencies too broad for a mechanical tracing pen must be displayed. Typically, the bandwidth that must be shown is from about 10 to 3000 hertz—a range readily handled by the oscilloscope.

We might expect that one or more electrodes placed in firm contact over a major muscle, such as an arm or thigh muscle, would yield a significant voltage that could be amplified, processed, and displayed. In practice, to obtain an accurate reading it



A. Normal EEG of a 49-year-old female.

B. Normal EEG of a 5-year-old female in an awake test.

Figure 8.14 Examples of EEGs (Tracings courtesy of Dr. John Mundall)

is usually necessary to insert at least one electrode—usually in the form of a very sharp, sterile needle—into the muscle itself. Measurements are then made between that electrode and a nearby reference electrode in contact with the patient's body. This type of connection is said to be *unipolar*. A bipolar approach may be more effective and requires the insertion of two electrodes into the muscle under study.

The amplified but unprocessed signal resembles a burst of noise made up of a number of electrical impulses of random frequency. After processing, it can be displayed on an oscilloscope with long-persistence trace on its screen. Signal levels run from perhaps 40 to 50 microvolts at the electrode source to nearly one millivolt (1/1000 volt). A built-in audio amplifier and speaker provides a characteristic crackling or rustling sound when the electrodes are properly inserted.

The EMG is principally used as a diagnostic instrument for nerve and muscle disorders. It is also helpful in evaluating a patient's progress in a program of medical rehabilitation.

Electrocardiograph

The **electrocardiograph** (EKG) is a medical-electronic instrument that amplifies and records the changes in electrical potential that occur during the beating of a human heart (Figure 8.15). Its primary use is providing information required in diagnosing heart action abnormalities.

Although Willem Einthoven, a native of the Netherlands, produced the first electrocardiograph about 1910, it remained for German medical technologists to improve it to a level of effectiveness where it would be generally accepted by the medical profession. The German spelling, *electrokardiograph,* gave rise to the acronym EKG.

The first EKG imported to the United States in the early 1920s was a massive piece of equipment weighing nearly 500 pounds. It was difficult to use and to maintain. It recorded on photographic film that had to be developed and processed before the trace could be read. Present-day models, in contrast, may weigh as little as 25 pounds. They are rugged, reliable, and portable.

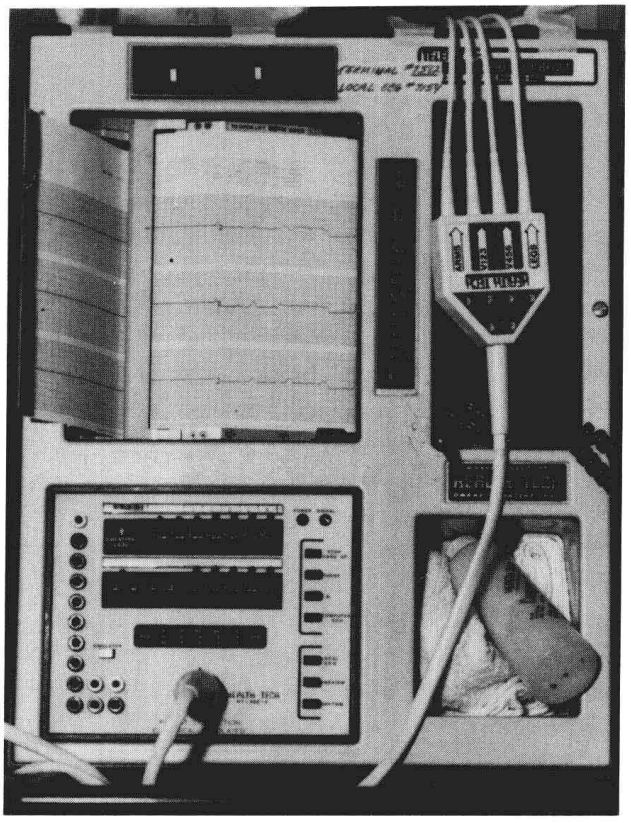


Figure 8.15 Portable EKG The electrocardiograph shown here is linked to a small computer that supervises its operation. Three tracing pens operate simultaneously from three output circuits. The machine can operate from both local and remote signal sources. (Photo courtesy of Sacred Heart General Hospital, Eugene, Oregon)

Einthoven's original placement of electrodes on the patient's body is still used today, although supplemented by additional placements and interconnections to provide further information. The usual procedure is to attach one sensing electrode to each wrist, one to the left leg, and one to the chest immediately over the heart. The fifth electrode, used

as a ground point in the system, is attached to the right leg. By switching within the machine, the various electrodes are temporarily connected in pairs. The effective voltage appearing on each paired electrode is compared to the voltage appearing on other paired electrodes. Paired electrodes are referred to as a *lead*.

When pairs of electrodes are interconnected, their effective voltage is what is actually being measured. The sequence of measurement is typically as follows: (1) left arm and right arm, (2) left leg and right arm, and (3) left leg and left arm. Since two electrodes are always interconnected, this type of recording is known as *bipolar*.

An additional approach that provides further data is measuring and recording the voltage on a single switch-selected electrode. The electrode sampled is called the *exploratory electrode*, and its potential is read against three interconnected limb electrodes that make up what is known as a central terminal, or reference point. The central terminal is assumed to have the same potential as the center of the body. EKG recordings made in this mode are referred to as *unipolar*.

As with the electroencephalograph, low-resistance contact between the silver electrodes and the patient's skin is essential since the voltages to be measured are so small. A small amount of a saline gel or paste is first applied to the electrode site to reduce the skin resistance. Then the electrode is strapped or clamped firmly in position over the treated skin.

The electronic amplifier unit used in the EKG machine must have high input resistance, high gain, and a very low level of hum and noise. Sufficient amplification must be provided to bring an electrode voltage measured in millivolts (1/1000 volt) up to a level sufficient to move the recording stylus across the moving chart. The recording mechanism is a type of D'Arsonval galvanometer using an electrically heated stylus that requires chemically treated paper.

By the application of a microcomputer to the EKG, all the various recording positions and modes can be carried on simultaneously to save time and cost. When this is done, a multiple-track recording is made on a single sheet of graph paper.

In addition to the usual graphical presentation on paper, the EKG output can also be used to produce an image on an oscilloscope screen. This technique has the advantage of high visibility, particularly if the trace on the instrument is set for maximum brilliancy. EKGs in this form find wide use in surgeries, emergency rooms, and recovery areas, where maximum visibility is important. However, for diagnostic work the chart form is still preferred.

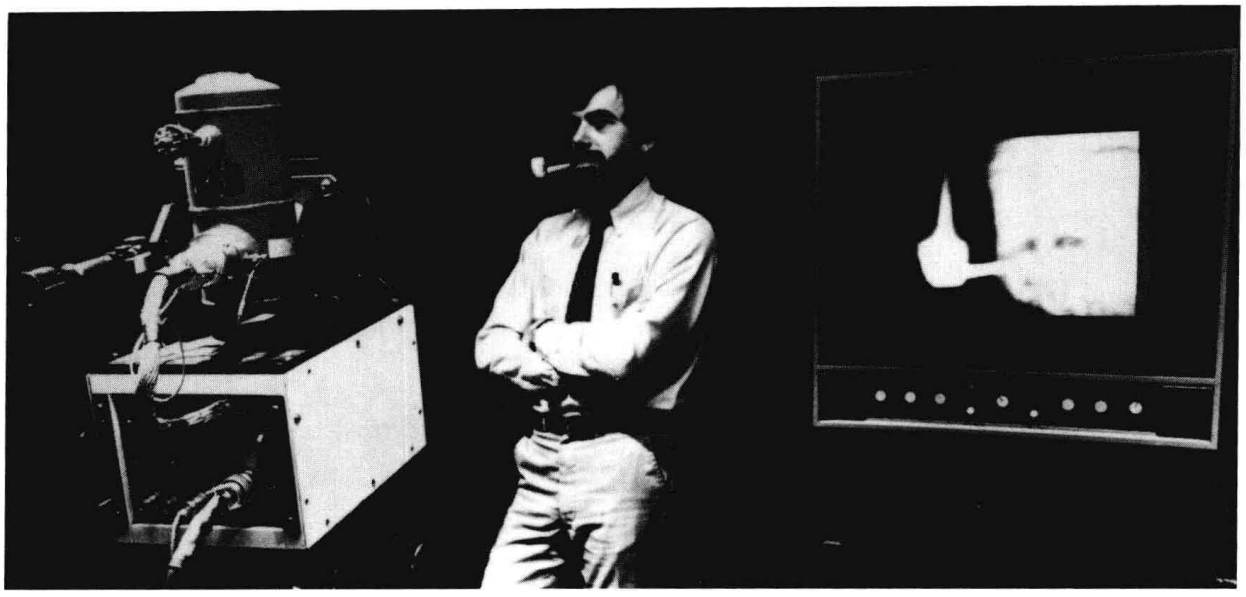
The EKG is probably the most widely used medical-electronic instrument in a typical hospital or healthcare unit. Its diagnostic value is so fully accepted that annual medical checkups are seldom considered complete unless an electrocardiogram has been made.

Thermographic Equipment

Thermography is a rapidly growing diagnostic technique that has proved particularly helpful in the diagnosis of breast cancer. It is based on the fact that malignant tissue exhibits a slightly greater temperature than that of the normal tissue surrounding it.

Since heat, like light, is an electromagnetic form of radiation, a TV camera and its associated video amplifier can be readily modified to operate on the nearby infrared frequencies and produce a display on a modified TV monitor. A malignant breast tumor will radiate more electromagnetic waves than nearby normal tissue. These waves will show up on the TV monitor's screen as an area different in color from the general background. This technology is effective only when the site under investigation is close to the surface of the skin so that significant radiation of heat waves can be detected. The equipment used (camera, amplifiers, and monitors) resembles a small, closed-circuit television installation.

While thermography by itself cannot positively diagnose cancer, it is very helpful for screening a large number of patients and indicating who should receive further tests. Thermography used to locate



RCA's experimental camera detects small variations in heat and creates recognizable TV pictures in total darkness. It uses a solid-state imager that contains over 8000 heat-sensitive elements. Its thermography application lies in its ability to display warm areas as white and cool areas as gray or black. (Photo courtesy of RCA)

potential breast cancer is usually known as **mammography.**

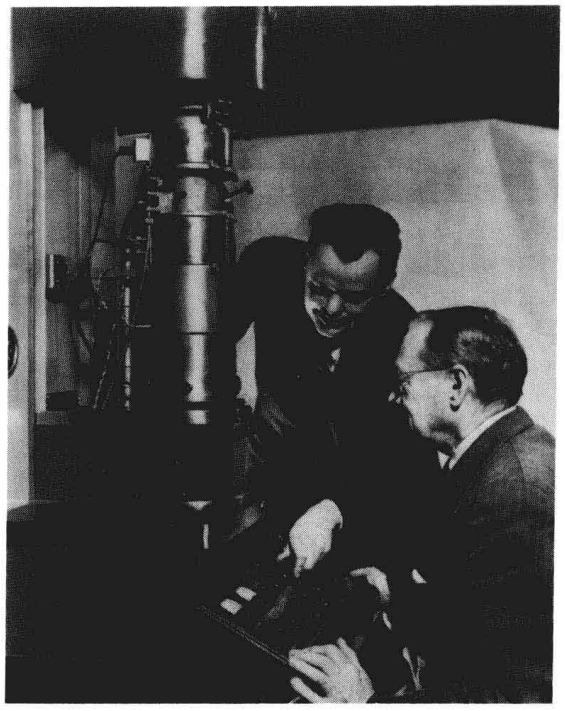
The high expectations originally held for thermography as a diagnostic tool have been realized only in part, and many of the procedures that planned to use thermography have instead been assigned to ultrasound. Nevertheless, thermography is very efficient in locating and interpreting circulatory problems in the head, the torso, the arms, and the legs.

Electron Microscope

One of the axioms of physical science is that no optical microscope can be constructed that will reveal any object whose dimensions are smaller than a wavelength of visible light. The upper limit of magnification for optical microscopes is currently a little over 2000 times. The electron microscope (Figure 8.16), making use of electromagnetic fields to focus clouds of electrons into a needle-sharp beam, can magnify an object perhaps 100,000 times—almost 50 times more than an optical microscope.

The first electron microscopes produced an image in much the same way as images were obtained by passing X-rays through an object and causing them to expose a photographic plate. The sharply focused electron beam was forced to pass through a thinly sliced piece of the material under investigation and then fall on a sensitized photographic plate below. The intensity of the beam was so great that recently living tissue or microorganisms were burned to a crisp, and only a shadow picture was obtained. (Living tissue and organisms are almost immediately killed by the high vacuum in the specimen chamber of the microscope.)

Borrowing from television technology, the newest electron microscopes scan the specimen under study. The needle-sharp electron beam moves both



The late Dr. Vladimir Zworykin (seated) and Dr. James Hillier test an improved version of the original electron microscope they developed in the early 1940s. (Photo courtesy of RCA)

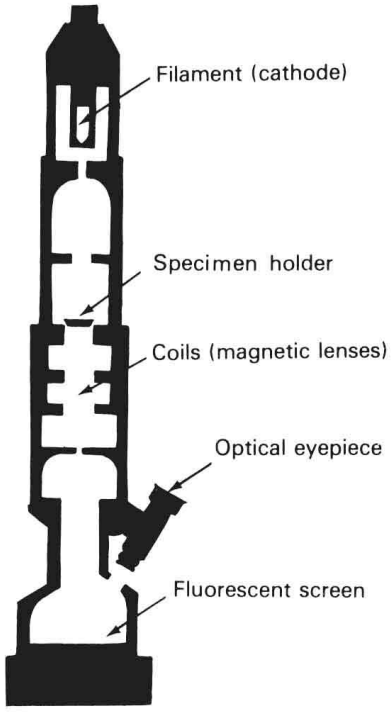


Figure 8.16 Cross Section of an Electron Microscope Although not clearly shown in drawing, a high vacuum is maintained within the microscope at all times.

from left to right (horizontal scan) and from top to bottom (vertical scan) with vastly improved results. The image is remarkably detailed, and the constant moving of the scanning beam reduces overall exposure so that recently living tissue and organisms can be studied at very high levels of magnification. For example, the pollen grain in the opening photo of this chapter was magnified 3000 times, yet the photo is remarkably detailed. Very high magnification levels also give a feeling of depth and perspective. Much of our knowledge of viruses and of the growth and development of various cell structures is traceable to the use of the electron microscope in diagnosis and research.

ELECTRONIC EQUIPMENT FOR THERAPY

Therapy is treatment to relieve pain or discomfort or to enhance a patient's sense of well-being. Since the late 1920s, a number of electronic machines have been developed for providing therapy.

Diathermy

Based on Arsene D'Arsonval's discovery that radio waves produce deep-seated warming in human tissue exposed to their effect, the first diathermy machines

This Raytheon microwave diathermy unit is used for deep therapy. The spacer keeps the electrode at a safe distance from the patient's wrist.

were essentially medium-power radio transmitters capable of generating 100 watts or so of radio frequency energy. Usually two 50 watt transmitting tubes were used in a balanced circuit that was quite inefficient, and the available output power was of the order of 100 watts. A frequency of approximately 27 megahertz was assigned to early diathermy units with the hope of minimizing interference with broadcasting and point-to-point communications. This same frequency is currently used for citizens band transmissions, and new diathermy equipment operates in the ultra-high-frequency (UHF) spectrum higher than 300 megahertz.

When power from the transmitter's output circuit is coupled to suitable coils or electrodes placed close to, say, the patient's aching shoulder or sprained ankle, dielectric losses caused by the radio waves passing through the body's tissue change most of the waves' energy into heat energy. This gentle but deep-seated heating may relieve pain and speed healing by improving blood circulation.

Hospital use of diathermy equipment seems to be slowly declining, probably because ultrasound therapy offers many of the advantages of diathermy and is easier to use. Also, ultrasound does not interfere with radio and TV reception—always a possibility when a radio transmitter is used. The use of diathermy continues in many smaller healthcare clinics and private offices. Newer diathermy units may be powered by a Magnetron tube (the same tube used in microwave ovens) and, although operating on the comparatively high frequency of 2500 megahertz, may generate well over 100 watts of radio frequency power.

Ultrasound

In diagnosis, carefully timed pulses of ultrasonic energy are generated electronically and fed to a crystal transducer that changes them into ultrasonic vibratory waves. As the pulse penetrates human tissue at a known rate of travel, a portion of it is always reflected to the source every time it encounters a change of density. The crystal transducer then transforms the returned echo back to an electrical signal that is amplified, processed, and applied to the screen of an oscilloscope for interpretation.

In the use of ultrasound as a therapeutic device, the electronic technology is markedly simpler because no receiving equipment is necessary to detect and amplify the echo pulse. Also, there is no need for an oscilloscope.

A transistorized oscillator and amplifier combination is required to generate ultrasonic energy in electrical form, and a husky transducer crystal is needed to convert it to a vibratory wave. Usually provision is made for supplying the energy delivered by the amplifier to the crystal in either a steady or a pulsed form. If pulsing is used, the frequency rate will be 1000 kilohertz or less. When supplying ultrasound continuously, the ultrasonic generator will probably have a rating of 15 watts or less; when pulsed, about 20 watts or less.

Because of the relatively large amount of power that must be handled by the transducer, a large, carefully laminated crystal of lead-zirconium titanate is used. Electrically, it closely resembles the ceramic cartridges used in crystal microphones and crystal stereo pickups.

When ultrasound energy is applied to the body's tissue by holding the crystal in contact with the skin, the effects are similar to those obtained by the application of diathermy. First, a gentle warming effect on the underlying tissue relieves the pain of sprains and muscle spasms. The warmth stimulates circulation, which speeds healing and tends to flush away harmful metabolic products associated with such injuries.

Ultrasound, as a vibratory form of energy, subjects both tissue and muscle fibers to **micromassage.** Because of the high frequency of the vibratory impulses, patients are not conscious of the massaging action. However, the relief from the soreness and stiffness that it brings is definitely noticeable.

Although not widely used by hospitals, ultrasound is in extensive therapeutic use in smaller healthcare centers and in private practice. Ultrasound holds promise as a therapeutic agent, particularly in the treatment of damaged or weakened muscles. Its effectiveness can be enhanced by using it in conjunction with electrical stimulation—that is, application through the use of metallic electrodes of moderate value of DC voltage normally pulsed at a varying frequency. The pulse normally has a very rapid rise and fall time—what the electronic technician calls a **square wave.**

Acupuncture

Acupuncture as presently used in the United States can be considered another application of electronics in medical therapy, despite the fact that some medical practioners would deny it professional status. In China, where the art of acupuncture was developed many centuries ago and has reached a high degree of sophistication, no external electrical source is normally used in treatment. The normal difference in potential between two or more points on the surface of the body is adequate for producing a weak flow of current along a nerve channel. In the United States, where acupuncture is a very new form of therapy, an external DC voltage source is almost invariably used.

Acupuncture's most widespread application in this country seems to be for pain relief without the use of drugs that might produce side effects or give rise to patient addiction. It is typically used in giving relief from migraine headache, low back pain, and arthritis.

In the treatment of migraine headache, for example, eight to ten very sharp, sterilized needles are inserted in the patient's body just below the surface of the skin. Typically, electrodes would be inserted at the following sites: in the neck at the base of the skull, behind each ear, over each eyebrow, at the crown of the head, and between each thumb and forefinger. The needles are interconnected by flexible, insulated leads and alligator clips and then connected to a battery-powered source that provides a pulsating direct current at a potential of perhaps 8 to 10 volts at a current strength of 200 microamperes (1/1,000,000 ampere).

The electronics of the control device can be extremely simple: Only one or two transistors and a few capacitors and fixed resistors are required. The principal adjustment is provided by a potentiometer (variable resistor), which controls the timing of the DC pulses, usually over a range from 175 to 200 pulses per minute.

The voltage supply must be adjustable over a moderate range to compensate for varying levels of body-tissue resistance. A 12 volt battery is often used for the DC source because of its convenience and as a safety precaution. (If an AC-powered DC power supply were used, a component failure could produce a fatal shock. A breakdown in the insulation between primary and secondary windings of the step-down power transformer could cause the 110 volt current to be applied to the electrodes and thus to the patient.)

A typical treatment lasts about twelve minutes and can be repeated in several days or spread out over a period of several weeks. Frequently three treatments may give nearly complete relief from migraine headaches. After a few months booster treatment may be necessary for maintenance.

Acupuncture practitioners differ as to why the treatment is effective in reducing local pain, but they seem to agree that acupuncture reduces the likeli-

hood of the transmission of pain sensations along a nerve channel that is electrochemical in its functioning. There may be other reasons for acupuncture's analgesic (pain-relieving) effect.

ELECTRONIC LIFE-SUPPORT EQUIPMENT

Certain medical-electronic equipment can support or in part take over a body function that is failing or likely to fail. Probably the smallest piece of electronic equipment in daily use in medical applications is the implantable **pacemaker,** or **heart pacer** (Figure 8.17). These units are designed to provide artificially the proper electrical impulses needed to initiate the cycles of the heart when the naturally occurring impulses are too slow, too weak, or irregular. By 1980 well over 200,000 patients were using pacemakers, and this figure is growing rapidly.

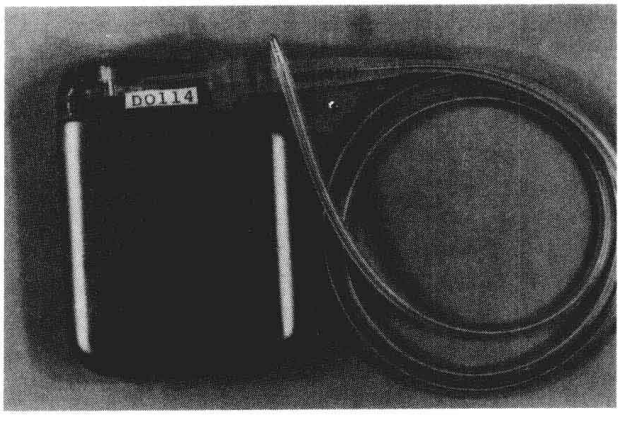


Figure 8.17 Implantable Heart Pacer Designed for surgical implantation on the patient's chest wall, this programmable pulse generator is powered by the new lithium-iodide batteries whose cell life is expected to be at least four years. The catheter with its enclosed platinum wire will be introduced into a vein leading to the heart. (Photo courtesy of Sacred Heart General Hospital, Eugene, Oregon)

Early pacemakers were so bulky that they were mounted externally, with two leads extending through the chest wall to the heart area to provide the controlling pulses. However, the point where the wires entered the body always posed the danger of infection and often made bathing and physical movement difficult. The newer heart pacers are remarkably compact and completely encapsulated in clear plastic. They can be safely implanted on the chest wall or elsewhere. With the new mercury-zinc batteries, cell replacement may not be needed for more than four years.

A pacemaker consists of three basic components: the batteries that serve as the power source, the electronic pulse-generating and sensing unit, and the electrodes that convey the electrical impulse to the heart. New, longer-lasting and more reliable batteries are constantly under investigation, and even miniature nuclear cells have been used as power sources.

The basic design of a heart pacer is quite simple. First, the pulse generator must be highly reliable, with a basic frequency usually variable over the range of 60 to 70 or 80 to 90 pulses per minute. Intensity of the pulse can also be varied. A sensing circuit is used to detect the electrical changes in the heart that precede a heartbeat. If a heartbeat fails to occur, the sensing unit decides whether to initiate its own pulse or to do nothing. A sophisticated unit of this type is referred to as a *demand pacemaker* (Figure 8.18). It prevents the problem that would occur if the normal rhythm of the heart were interrupted by an unwanted pulse or series of pulses from the pacemaker.

Typically, the pacer capsule is implanted near the right armpit. The two platinum or iridium electrodes operate in a unipolar mode and are dealt with separately. The positive lead (electrode) is merely implanted in nearby tissue. The negative lead is threaded into a sterile catheter of small diameter that the surgeon inserts into a vein leading to the heart. The catheter's progress can be observed on a fluoroscope. Installation is accomplished with a min-

Figure 8.18 Portable Demand Pacemaker This type of heart pacer is suitable for use in any hospital location. It uses standard alkaline D cells as a power source and is easily maintained. (Photo courtesy of Sacred Heart General Hospital, Eugene, Oregon)

imum of danger, and recovery from the procedure is usually rapid.

MAINTAINING MEDICAL-ELECTRONIC EQUIPMENT

Routine maintenance can be performed on medical-electronic equipment by the health professional assigned to its operation. Little or no knowledge of electronics is required. In many cases equipment malfunction is caused by simple problems, such as plugs and sockets making poor connection to cables, circuit breakers tripped, fuses blown, and, not infrequently, an AC power cord not plugged into the outlet.

Maintenance can be expedited if a log of the daily meter readings on the instrument's control panel is kept. On many pieces of equipment, meters are provided to read voltage and current values. Deviation from normal values may provide clues to the location of a problem. A record of meter readings that shows a gradual, day-to-day change in current or voltage values may signal the need for some preventive maintenance.

Usually, a technician who can handle the complex circuitry of a color TV set and can interpret common waveforms on the screen of an oscilloscope can successfully train on the job as a biomedical equipment technician (BMET). This title implies that the technician's knowledge transcends electronics alone.

Should a power transformer break down or a capacitor short circuit, the patient's only protection from shock or possible electrocution may be the solidness of the grounding system used to interconnect all pieces of medical equipment. Hospitals always have several instruments known as *ground-fault locators* that can quickly reveal the lack of a ground connection or the presence of a high-resistance ground, which may be equally dangerous. The machine operator and the maintenance technician are responsible for maintaining the integrity of all ground connections at all times. The operator should never disconnect equipment and then reconnect it without a ground check by the technician. Also, the operator should never endanger the patient through the use of adapters that make it possible to plug an instrument with a 3 pin power cord into a 2 pin wall receptacle. The third pin (round) is the ground connection, and it must be firmly connected to the ground system. New standards for medical-electronic equipment, in effect in 1983, require that all patient connections be isolated from ground.

SUMMARY

Although the early work of Wilhelm Roentgen with X-rays and of the Curie family with radioactivity is well documented, the contributions of Arsene D'Arsonval, who pioneered the use of electronics in medicine, are seldom recognized. Prior to 1910, D'Arsonval made significant discoveries in the application of electronics to medicine in such areas as the use of concentrated radio waves for deep-tissue therapy and the use of Bell's telephone as an amplifier for the weak voltage changes in human tissue caused by neural action. These two applications led to the development of diathermy and of the electrocardiograph. The tracing pens used in modern analytical equipment, such as the EKG, the EMG, and the EEG, are based on the D'Arsonval movement. D'Arsonval was also the first to realize the urgency of immediate artificial respiration for victims of electrical shock.

In the last fifty years, and largely within the last three decades, electronics has revolutionized the medical profession's approach to diagnosis, therapy, and patient care. Much progress has been made possible by the development of highly specialized transducers that convert blood pressure, temperature, and pulse and respiration rates into weak electrical impulses. After amplification by active devices and processing by electronic circuits, these impulses are applied to sophisticated display devices such as oscilloscopes or TV monitors. The same impulses can also be used to actuate tracing pens that move across a moving strip of graph paper and provide a permanent record. The most recent and most sophisticated medical-electronic machines employ a computer to organize sensor-recovered data for presentation on the video screen. Computer axial tomography (CAT) scanners and nuclear magnetic resonance (NMR) scanners could not function without an integral and specially designed computer.

Electronics is also being employed in therapy, treatment to relieve pain or discomfort. For example, expanded use of ultrasound provides both the deep and superficial micromassage to ease pain and speed healing without side effects.

Certain electronic equipment can support or in part take over a body function that is failing or likely to fail. The best example is the surgically implanted cardiac pacemaker. Completely encapsulated in plastic, active devices electronically generate precisely timed pulses of DC voltage that keep an ailing heart beating regularly. Pacemakers have increased the life expectancy of thousands of patients with cardiac problems.

A well-trained electronic technician should experience little difficulty in the repair and maintenance of medical-electronic equipment. The components and circuitry are the same as those used in other electronic fields. Even the test equipment used in servicing is the same. One important factor must be constantly stressed: keeping patients safe from electrical shock. All electronic equipment must be securely grounded to the healthcare facility's ground system and must be regularly checked with a special ground-fault locator.

KEY WORDS AND CONCEPTS

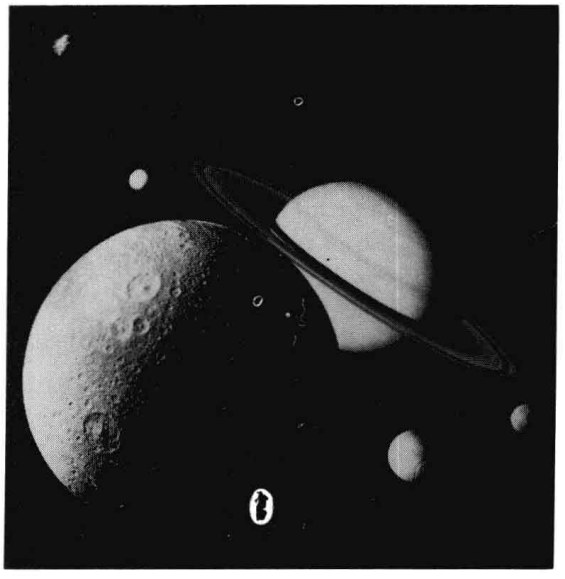
acupuncture
biophysics
cathode-ray tube
CAT scanner
diathermy
echocardiography
electrocardiograph (EKG)
electroencephalograph (EEG)
electromyograph (EMG)
fluoroscopy
heart pacer (pacemaker)
laser
luminescence
mammography
micromassage
NMR scanner
pulse-echo diagnostics
scanning
scintillation detector
therapy
thermistor
thermography
tomography
square wave
ultrasound
X-ray

SELF-TEST QUESTIONS

1. What diagnostic device of modern medicine did Roentgen discover in his experiments with electronics?
2. What D'Arsonval invention is widely used in medical-electronic instruments that record data on a moving strip of graph paper?
3. Why are oscilloscopes used extensively to display critical functions such as heart action, respiration rate, and temperature?
4. How are low-power portable radio transmitters used in patient monitoring in some coronary care facilities?
5. In your own words, describe how an accelerated beam of electrons produces high-energy X-rays.
6. What was Edison's principal contribution to medical electronics?
7. What is the advantage of using a CAT scan rather than conventional X-rays for patient diagnosis?
8. When ultrasound is used as a diagnostic tool, a special piezoelectric crystal serves as a transducer with two functions. What are those functions?
9. What types of brain damage are readily identified through the use of the electroencephalograph?
10. What is the function of the conventional audio amplifier and loudspeaker incorporated in the electromyograph?
11. What advantage is there in controlling the operation of an electrocardiograph with a microcomputer?
12. Thermography as a diagnostic tool is based on its ability to detect very small variations in the amount of heat radiated from the patient's skin. Of what significance is an area of slightly elevated temperature?
13. Why do newer types of scanning electron microscopes permit the viewing of organisms and tissue that older, nonscanning microscopes instantly destroyed?
14. In medicine and the health occupations, what is meant by therapy?
15. Why has diathermy caused extensive interference with radio reception?
16. Why is it simpler to design and construct ultrasonic equipment for therapeutic use than for diagnostic purposes?

17. Why do acupuncture specialists prefer to use batteries as a DC power source rather than a power supply connected to the 110 volt line?
18. Why must a perfectly functioning implanted heart pacer be periodically removed from the patient's chest, serviced, and reinstalled?
19. How can the daily log of voltage and current readings of a piece of medical-electronic equipment enable the maintenance technician to make prompt repairs?
20. Why is it important that all medical-electronic equipment be solidly connected to a dependable grounding system?

9 Electronics in Space and World Exploration



This montage shows the cratered surface of moon Dione (foreground) with Saturn and four more of its moons in background. The photo, taken by Voyager I in November 1980, illustrates the remarkable detail rendered when the image of an all-electronic camera is computer enhanced. (Photo courtesy of NASA)

Historical Perspective

An Overview of NASA Electronics
- Basic Requirements for Space-Vehicle Electronics
- Basic Requirements for Ground-Station Electronics
- The Space Program's Effect on Electronics

Highlights of NASA's Space Program
- Early Period of Space Exploration
- Space Probes
- Space Shuttle Orbiters

NOAA
- NESDIS
- Weather and Environmental Satellites
- Landsat and the Multispectral Scanner

Summary

Key Words and Concepts

Self-Test Questions

In this chapter, we describe how two of the largest public agencies—the National Aeronautics and Space Administration (NASA) and the National Oceanic and Atmospheric Administration (NOAA) use electronics. Both agencies use advanced electronics in daily and long-term operations. Few of their programs could function without electronic equipment. The impact of these agencies' multibillion dollar budgets on the Electronics Industry has been reflected in the industry's direction for the past two decades.

HISTORICAL PERSPECTIVE

In October 1957, the Soviets launched the world's first artificial earth satellite. Known as Sputnik I, it was basically an aluminum sphere about the size of a basketball and with rod-shaped antennas attached. Its electronics were of the most elementary form: A low-powered transmitter aboard constantly repeated a single, brief message in international Morse code. As Sputnik tumbled through the night sky, it gave off flashes of reflected sunlight that served to remind watchers that the United States, long acknowledged as a world leader in the field of science, was well behind the Russians in space technology.

Congress soon learned constituents felt strongly that America should match the Soviet accomplishment with a satellite of its own. As pressures mounted, President Kennedy, who had taken office in 1960, announced his belief that the nation should commit itself, before the decade was out, to landing on the moon.

Congress was quick to respond. In 1958 an act creating the National Aeronautics and Space Administration passed both houses with scarcely a dissenting vote. The financial support provided was practically unlimited. Although NASA was to have responsibilities in the aeronautics field, Congress made it clear that the greatest priority was the space program, which not only included a landing on the moon but exploration of interplanetary space and even of the planets themselves.

By 1965 NASA's annual appropriation had reached $5.2 billion, and during the Apollo program 200,000 persons were directly or indirectly involved in the program. This commitment had an enormous impact on the economy of the Electronics Industry.

AN OVERVIEW OF NASA ELECTRONICS

Basic Requirements for Space-Vehicle Electronics

The basic requirement for any type of space vehicle is a completely reliable, low-voltage, high-current power supply—invariably a battery of special design kept fully charged at all times by hundreds of silicon solar cells (connected in both series and parallel to provide the right charging voltage). These cells are mounted on the outside of the space vehicle, either on its surface or on connected panels. This basic DC power supply operates all transmitters and receivers used for two-way communication and runs all sensors providing electronic data on temperature, electromagnetic radiation encountered, and light spectrum measurements. All computers and microprocessors aboard are also powered from this primary source.

The more sophisticated space vehicles are designed to carry one or more astronauts. These vehicles require additional sensors and the supplementary equipment needed to transmit their electronic output signals in digital form to the ground station and its associated command center. The data transmitted always include information on pulse and respiration rates, blood pressure, and as many additional measurements as the medical team at the ground command station feels are relevant. Sensors provide a constant stream of data flowing earthward and delivering information on the crew's basic functions. Additional sensors may sound warning buzzers or light flashing lamps within the spacecraft to alert the astronauts of such dangers as depleted oxygen levels or excessive accumulations of carbon dioxide. In

certain cases sensors may electronically control the action of microprocessors and related circuitry to open valves or operate systems to correct improper conditions.

Communicating with the Ground Station

The critical link between the space vehicle and the ground station is the system of antennas mounted on the spacecraft's exterior and connected to the inboard transmitters and receivers. Because of the very high frequencies and the very short wavelengths used by NASA for communications (usually somewhere between 2 to 100 gigahertz), antennas can be very small physically and still maintain high efficiency. A gigahertz (GHz) is 1000 megahertz (MHz).

A rather high degree of directionality is desirable in the spacecraft's antenna system because of the limitation on the amount of transmitter power available. This limitation is imposed by the permissible drain on the craft's primary battery, which must be kept charged by the output of silicon solar cells mounted on the space vehicle's exterior. Directional antennas can be compared to a flashlight beam. If a conventional, two-cell flashlight is taken apart and the unit reassembled without the concave reflector behind the bulb, the intensity of light given off appears quite insignificant because its radiation is dispersed over a wide area. Reassembled with the concave reflector, the flashlight gives off a beam of high intensity and brilliance because the reflector concentrates and focuses the radiation into a very narrow path. Similarly, a spacecraft's antenna will probably use a **parabolic reflector,** or dish, to concentrate the transmitter's comparatively low output power into a signal beamed to the ground station.

Engineers have long been aware that an antenna system that will transmit well in a certain direction will also serve admirably to receive signals from the same direction. The system exhibits about the same gain in the receive as in the transmit position. The same antenna, then, is very commonly used to enhance the strength of both incoming and outgoing signals. This technique reduces the number of antennas required by each space vehicle and simplifies the exterior mounting of antennas that are required. The principal disadvantage of a highly directional antenna system is that it must be aimed at the ground station with considerable accuracy. An antenna 30° or 40° off target may prove nearly worthless for communication.

NASA engineers still prefer to use vacuum tubes rather than transistors in the final stages of their space-vehicle transmitters. Output power is usually within the range of 10 to 100 watts. While transistors are considered acceptable for power values of 25 watts output or less, NASA engineers insist that vacuum tubes be used for output powers of more than 25 watts. Their thinking is that tubes can produce higher power at these frequencies (in the gigahertz), are more reliable, and have a longer life span than transistors. Special vacuum tubes called *klystrons* and *traveling-wave tubes* are used in these applications.

Basic Requirements for Ground-Station Electronics

The most important piece of equipment at the ground station, which intercepts the radio signals from the space vehicle, is an extremely high gain and highly directional receiving and transmitting antenna. The high gain and directionality are achieved by using a very large parabolic reflector, or dish. You've probably seen similar, but much smaller, installations on the roof of your local telephone company. For receiving purposes, the incoming signal is picked up at the focal point of the parabola and sent to the receiver circuits. For transmitting, the transmitter output is fed to the same focal point. In general, there is reciprocity between both receiving gain and directivity and transmitting gain and directivity. The gain is essentially the same for both functions.

Physical limitations determine how large a parabola can be rotated or positioned with accuracy in a high wind. Generally, bigger is better within those limitations. The largest parabolic antenna operated by NASA in the continental United States is at Gold-

stone, California, in the Mohave Desert just east of Los Angeles (Figure 9.1). The station, which has a parabola 64 meters (about 221 feet) in diameter, serves as a key station of NASA's deep space network. Transmission of instructions and response to the command can be accomplished between the Goldstone station and a spacecraft 625 million miles away in just 2 hours and 21 minutes.

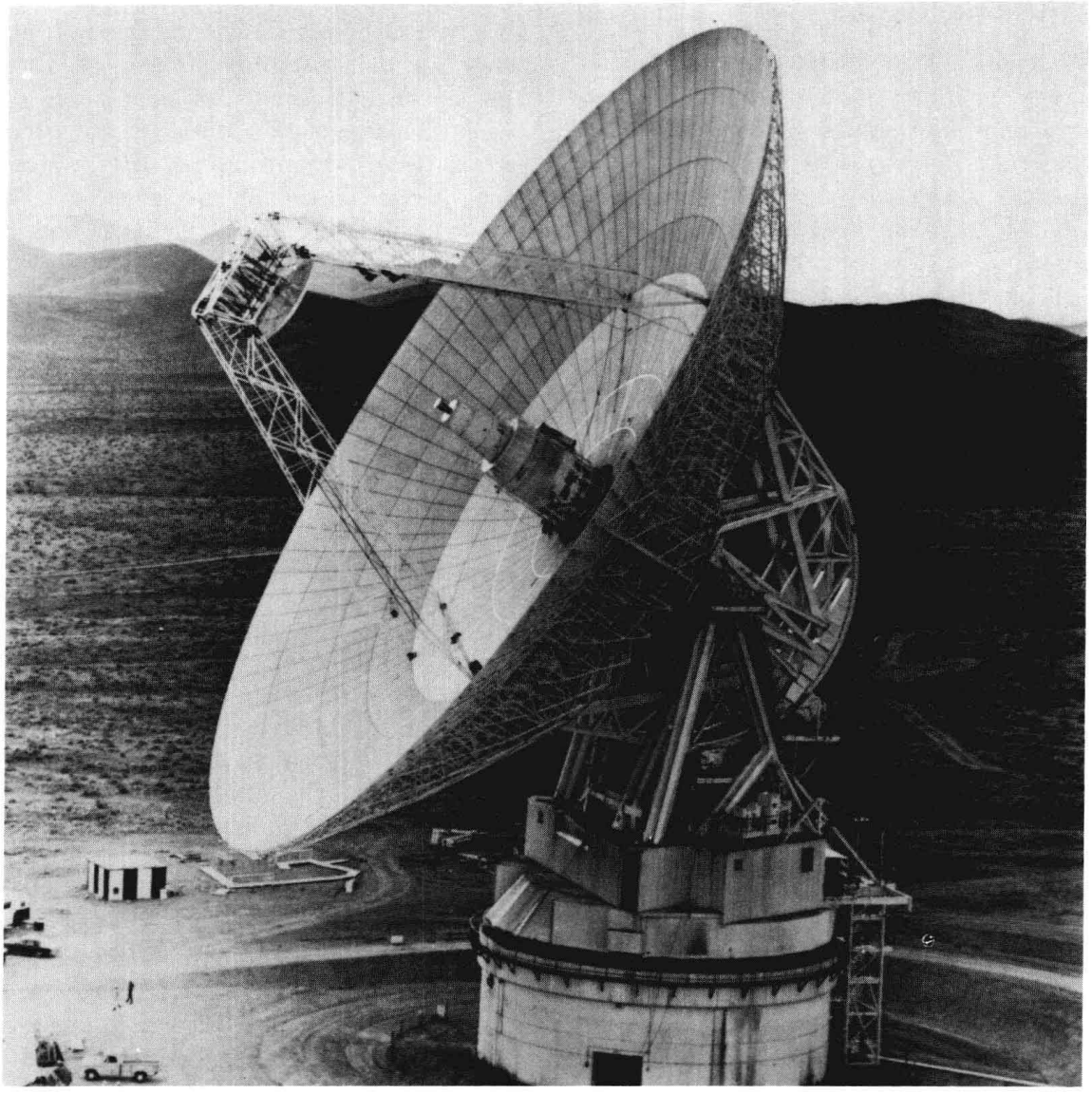


Figure 9.1 Parabolic Antenna Located at Goldstone, California, this antenna is used for both reception of satellite signals and transmission of electronic commands to the satellites. Its great size provides the high degree of sensitivity required for the reception of weak signals from millions of miles out in space. (Photo courtesy of NASA)

Much of the electronic gear in the operations room at the Goldstone site (Figure 9.2) would be functionally worthless if a poor or mediocre antenna installation were in use. The electronic equipment required at the ground station falls into five general categories: (1) receivers used for spaceship communications; (2) transmitters for the same purpose; (3) monitoring equipment, largely visual, for determining proper operations of receivers and transmitters; (4) data storage equipment for digital information from the spacecraft; and (5) equipment for transmitting either stored or directly received data to the mission operations center. In the case of manned space flights, the center was at Houston, Texas. For the current series of space probes, the Jet Propulsion Laboratory at Pasadena, California, assumes this role.

Because the earth revolves on its axis once every twenty-four hours, a whole series of ground stations, properly positioned, is required to maintain continuous communications with the spacecraft. Ground stations have therefore been constructed on leased sites on every continent. The data are relayed back to the command center by geosynchronous communications satellites stationed over the same spot on the earth's equator at a height of about 23,000 miles. Less critical data can be transmitted over conventional shortwave radio channels.

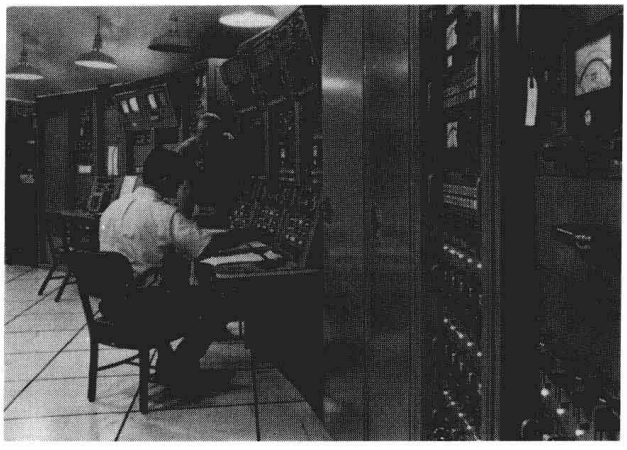


Figure 9.2 Operations Room for Tracking and Communications at Goldstone, California Electronic equipment includes sophisticated receivers and transmitters capable of operating on many communications channels. The Mohave Desert location is almost completely free from man-made electrical interference. (Photo courtesy of NASA)

The Space Program's Effect on Electronics

By far the greatest demand of NASA's space program was a reduction in the size and weight of electronic equipment and modules. The reduction was dictated almost completely at first by the comparatively low thrust of the early rocket motors used for launching. Progress was slow at first, but as high-density circuit boards and modular construction with tightly packed components developed, significant reduction in both weight and size of computers, transmitters, and receivers was achieved. By the mid-1960s Texas Instruments, Fairchild, Motorola, and RCA were involved in the development of the highly innovative integrated circuit—a unit with circuits of great complexity mounted on a tiny silicon chip. By 1970 Ted Hoff of Intel had developed an integrated circuit that contained the entire central processing unit of a computer on a chip less than 1/4 inch square and of inconsequential weight.

As electronic units shrank in size and weight, NASA engineers could cram additional sensors, receivers, computers, and even TV cameras aboard spacecraft. By ingenious switching systems, defective units could be switched out of the circuit and their backup replacements immediately switched in. Space travel became immeasurably safer for astronauts as the ultimate tragedy of space flight—a vehicle circling forever beyond the aid of the command center because of a massive failure of electronic control units—became less likely. The trend toward equipment redundancy was further aided by improvements in rocket motors that made possible the launching of far heavier space vehicles.

Closely related to the safety factors afforded by the backup concept of electronic equipment was NASA

engineers' continuing insistence on maximum possible reliability, whatever the cost. This consideration further increased the priority afforded NASA over private equipment purchasers in the marketplace. While insistence on complete reliability had a long-run beneficial effect on all electronics users, NASA's near-complete economic dominance of the Electronics Industry over a period of several years was to prove disastrous when the Apollo program was terminated in 1972 and NASA spending sharply cut back. Thousands of highly skilled technicians were laid off, while senior engineers and technicians accepted pay cuts and demotions in rank. Today the industry places more reliance on the private sector of the market.

HIGHLIGHTS OF NASA'S SPACE PROGRAM

For many years NASA has worked closely with government and private agencies in the design of suitable satellites for their particular needs. NASA conducted the launch of these satellites into orbit and assisted the sponsoring agency with positioning the satellite in space and maintaining a constant flow of radio communications between the satellite and ground stations. Since NASA already had a highly trained staff of aerospace and electronics engineers and several well-equipped launch sites, use of NASA's facilities avoided the need for duplication. The result was speeded development and launching of highly specialized satellites and great savings to both the government and the private sector.

Since nearly every NASA launch has been a news event, this section provides an overview of the earliest missions and then focuses on the current series of space-probe missions to Mars, Venus, and Jupiter and on specific projects like Spacelab and the space shuttle. Also discussed are the Voyager missions to Jupiter, Saturn, Uranus, and Neptune (not to be completed until 1989).

Early Period of Space Exploration

In 1962 John H. Glenn made the first significant space flight. He completed three orbits of the earth before splashdown at sea. All electronic systems worked perfectly and encouraged a more ambitious manned satellite series known as Gemini (twins), which would carry two astronauts aloft to maintain a twenty-four watch over the scientific and control equipment aboard. Gemini was a proving ground for the technical equipment and techniques that would be used in the Apollo program, which had as its goal President Kennedy's proposal to land an astronaut on the moon.

Only after six unmanned flights proved the reliability of the new Saturn V launch rockets was the first manned flight with three astronauts attempted. Three of these flights were in the vicinity of the moon, but with no landing attempted. Electronic communications proved faultless.

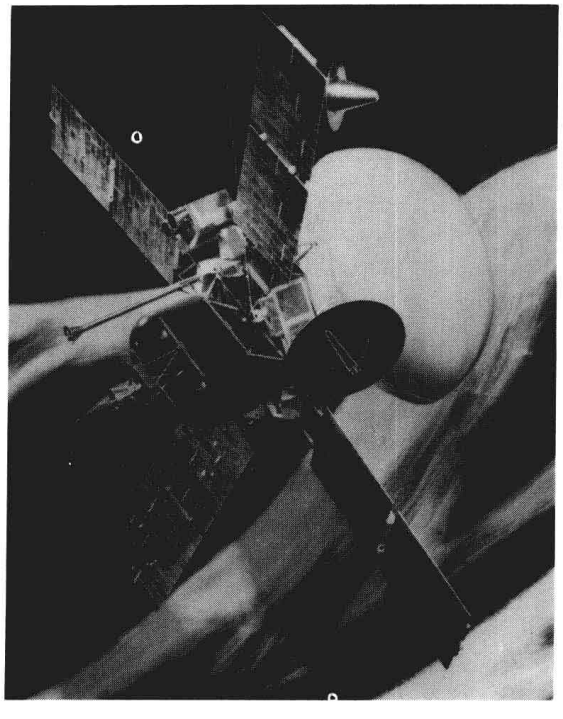
When Apollo XI was checked out (16 July 1969), NASA decided all systems were "go," and the mission proceeded exactly as planned. Using the lander, astronauts Armstrong and Aldrin descended to the moon's surface while Collins piloted the orbiting spacecraft. Armstrong's climb down the steel ladder to the moon's surface was the greatest news event in the history of television.

Six subsequent Apollo missions to the moon each had a different destination. Five were successful. One had to be aborted because of the rupture of an oxygen tank in midflight. Again, all electronic systems worked flawlessly, and the only electronic failures were those caused by the astronauts themselves. For example, a TV camera was ruined when allowed to point for seconds at the brilliant disc of the sun. The final Apollo mission concluded without incident in 1972.

Space Probes

At the conclusion of the Apollo program, NASA turned to the use of **space probes** for further explo-

ration of the planets and interplanetary space. Space probes are fully automated and highly specialized spacecraft equipped with a wide spectrum of sensors, including a number of TV cameras. Probes also carry laboratory equipment to conduct measurements of atmospheric gases and moisture content and to detect the presence of primitive life forms. Most of the earlier probes were of the "fly-by" type—for example, the probes that obtained close-up views of Mars, Jupiter, Venus, and Mercury without landing. In 1975, however, the space probe program was expanded to include planet orbiters and landers. The Voyager missions to Jupiter, Saturn, Uranus, and Neptune, started in 1977, will not be completed until 1989.



This full-scale engineering model of the Viking spacecraft is seen against a "Marscape." Space cameras, parabolic antennas, and solar panels for providing electrical power can all be seen on the model. The Viking mission was the first to place fully equipped laboratories on the surface of Mars while orbiting units overhead relayed data back to Earth. (Photo courtesy of NASA)

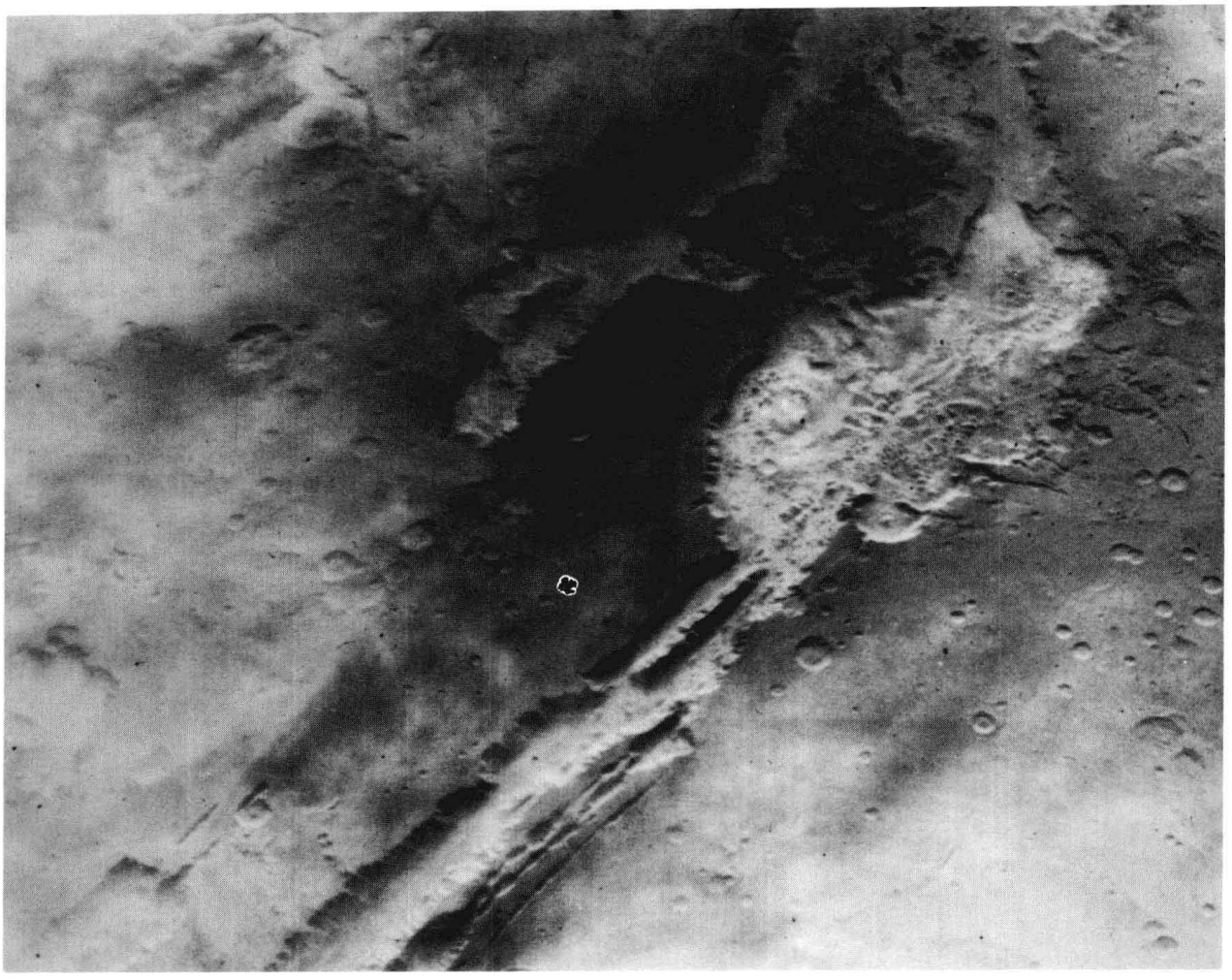
The Viking Mission

The Viking mission of 1975 was unusually ambitious in that it sent two orbiters with their landers to separate sites on Mars, some 35 million miles from Earth. Since then, both orbiters and landers have produced an enormous amount of information about the red planet. Some of the most compelling has been in the form of beautifully detailed aerial photos taken by TV cameras mounted on the orbiting units. Thanks to computer enhancement, the photos show impressive views of the waxing and waning polar icecaps; enormous rift valleys, apparently partly carved by water; and Olympus Mons, a volcanic mountain nearly three times as high as Mount Everest. Although most of the equipment used in the Viking mission was projected to have a life expectancy of perhaps two years, much of it was still operational years later, although constantly subjected to extremes of cold and heat.

Pioneer Venus I and II

A mission consisting of one orbiter known as Pioneer Venus I and five atmospheric probes known as Pioneer Venus II arrived in the vicinity of that cloud-hidden planet in December 1978. The orbiter was sent on an eight-month orbit of Venus to make measurements and take pictures, while the instrument-laden probes of Pioneer II descended through the super-heated Venusian atmosphere to the surface. They were all destroyed by the great heat within a short period of time, but not before valuable data were relayed to the orbiter.

Data from the orbiter's first map, based on radar soundings of the surface of Venus, suggest that Venus's topography could be similar to Earth's, with high mountains and extensive, relatively flat areas,



From a height of 10,000 miles above the surface, Mars's Grand Canyon is impressive even through haze. Geologically, it resembles Earth's Rift Valley in eastern Africa. Area shown is about as large as the state of Alaska. (Photo courtesy of NASA)

although there are no large bodies of water. Other instrumentation shows that the clouds are primarily composed of oxygen, water vapor, and sulfur dioxide.

Voyager Missions I and II

The Voyager missions, the most ambitious of the unmanned probe series, are still incomplete. The goals of the mission to the planet Jupiter were to study its composition, winds, and magnetic field; to search for more Jovian moons (Galileo discovered four moons in the 17th century); and to determine the moons' temperatures, atmospheres, and physical composition.

On the mission to the planet Saturn, the goals were much the same but also included a thorough study of the provocative rings that have puzzled scientists since Galileo's first crude telescope revealed them. After studying Saturn, Voyager I's mission

According to the planned trajectory of space probes Voyager I and II, Voyager I is now headed for interstellar space. Voyager II will fly by Uranus in 1986 and Neptune in 1989 before entering outer space. Note that the dish antenna must always be sharply focused on Earth. (Photo courtesy of NASA)

was complete, and it started on a course that would lead it forever into the cold, empty darkness of interstellar space. Voyager II has been programmed to rendezvous with Uranus in 1986 and Neptune in 1989 and to provide scientists with the first detailed information on these planets since their discovery.

Neither of the Voyager missions could have been accomplished without the most sophisticated electronics both in the spacecraft and on the ground. Course corrections and commands to the cameras and other sensors were sent from the earth by radio waves bearing encoded digital signals. At the same time the spacecraft transmitted a flow of information, again digitally encoded, back to the best-located ground station. Although radio waves, like all electromagnetic waves, travel through space at the rate of nearly 186,300 miles per second, the distances in interplanetary space are so vast that half an hour or more may elapse before a transmitted signal is received. Transmission from Earth to Jupiter, for example, takes 37 minutes.

Equipment and Orbits. Although both the Voyager I and Voyager II space probes were identical in appearance, minor differences existed in their instrumentation. The major difference was that Voyager II had superior cameras from both an electronic and an optical standpoint (Figure 9.3). Actually, Voyager II was considered a backup probe for Voyager I. In practice, Voyager II's superior camera resolution (rendering of fine detail) discovered noth-

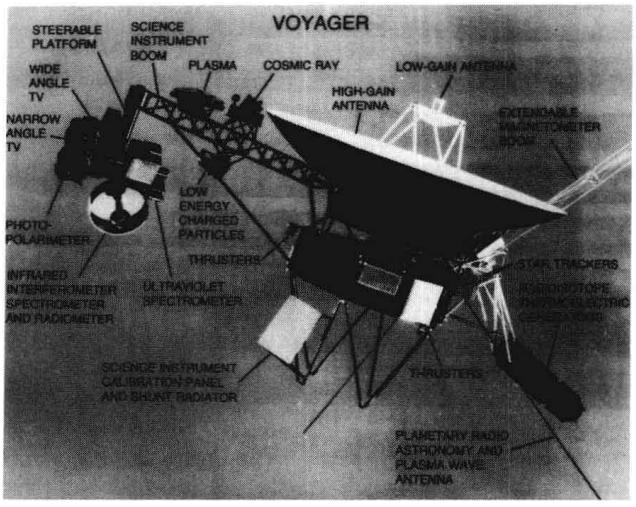
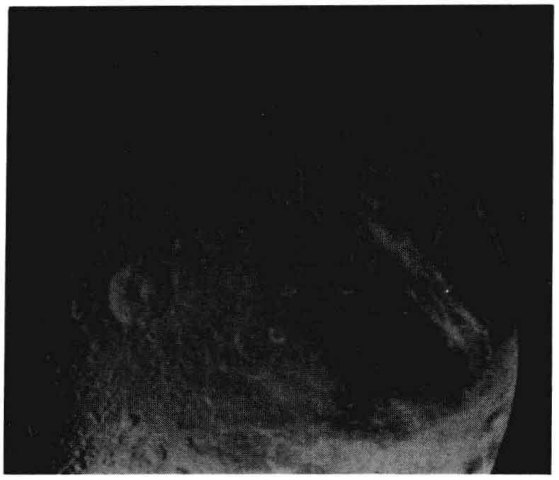


Figure 9.3 Voyager Space Probes Equipment Voyager I and Voyager II are essentially identical in appearance, but Voyager II has superior cameras. (Photo courtesy of NASA)



A space probe's all-electronic cameras, digitally enhanced, can produce a picture of remarkable detail. Shown is one of Saturn's moons, Dione, as seen from Voyager I at a distance of 150,000 miles. Its surface pocked with impact craters, it closely resembles Earth's moon but is less than 700 miles in diameter. (Photo courtesy of NASA)

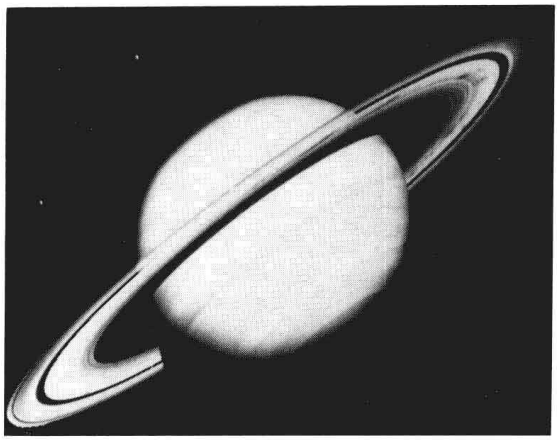
ing that the cameras of Voyager I had not already disclosed, but scientists welcomed the greatly increased detail. (Output of the probe's main transmitter was 8 watts.)

Launched two weeks before Voyager II in 1977, Voyager I was in the faster orbit and arrived at Jupiter on 5 March 1979 and at Saturn on 12 November 1980. Voyager II, on a much longer route, arrived at Jupiter on 5 July 1979 and at Saturn on 26 August 1981. Voyager II will arrive at Uranus in 1986 and at Neptune in 1989.

Mission Highlights. Some highlights of the Jupiter fly-by were the discovery of volcanic activity on the moon Io; the disclosure in photographs of a small, Saturnlike ring surrounding the planet as well as a small, potato-shaped moon—its sixteenth; and a highly detailed look at Jupiter's Great Red Spot, which proved to be a persistent and turbulent high-pressure area in the planet's atmosphere.

On the Saturn fly-by, the planet's spectacular rings were naturally the focus of greatest attention. Most scientists theorize, after careful study of the digitally enhanced photos, that the rings are made up of crystals of water-ice that have accreted (grown together) to form particles perhaps a meter in diameter. In bands of the rings where the particles are abundant, sunlight is brilliantly reflected. Where other bands of rings appear less brilliant or even dark, the particles are either less abundant or very few in number. Two tiny moons were discovered embedded in the rings, and apparently a number of tiny moonlets are also enmeshed in the rings. Seven more moons were found, for a new total of seventeen.

Analysis of the printouts of the many electronic sensors aboard both spacecraft has led scientists to conclude that Saturn, like Jupiter, is basically a ball of hydrogen and helium without a firm surface and with winds of enormous velocities that create constant turbulence. The beautifully detailed images of Saturn's moons, like those of Jupiter's, have raised questions of origin that will require decades to answer. Most of Jupiter's moons and nearly all of Saturn's are smaller than Earth's moon.



Saturn, its rings, and 2 of its 17 moons were photographed by the cameras of Voyager II at a distance of 21 million miles. The rings are thought to be composed of meter-sized lumps of gritty water-ice. The darker bands represent thin areas where fewer particles reflect the sunlight. At lower left, the planet's shadow truncates the rings. (Photo courtesy of NASA)

Problems for Voyager. As mentioned earlier, Voyager II has been programmed on a course that will take it to Uranus by January 1986 and to even more obscure Neptune by August 1989. Much later, Voyager II will cross the orbit of tiny Pluto and enter interstellar space.

However, certain problems are still to be overcome if Voyager II is to accomplish all its objectives. The most pressing problem involves the reliability of radio communications at extreme distances. Radio waves (and all electromagnetic waves) exhibit a decrease in intensity as they leave their source, and the decrease varies inversely as the square of the distance. (At twice the distance, the field strength has fallen to one-fourth, for example.) The question arises, then, whether Voyager II will obey commands from the ground to perform vital functions like course correction, maneuvering cameras into proper position, or activating the various sensors. (Even nearing Saturn, Voyager II exhibited some problems.) The communications problem could be reduced by increasing the intensity of the signal received from the ground station through an increase in transmitter power, in the gain of the transmitting antenna, or in both. However, it seems unlikely that these measures will be taken because too much time would be needed for budgeting, planning transmitter and antenna changes, and actual construction.

Another problem for Voyager II involves electrical power. All the space probe's electrical power, essential for the operation of every servomechanism and electronic device aboard, is derived from a storage battery kept continuously charged by panels of silicon photovoltaic cells on panels exposed to sunlight. As the distance of the spacecraft from the sun increases, the cells' effectiveness decreases because the sun's radiation intensity fades. Therefore, electrical energy must be hoarded by turning off all unnecessary energy-consuming electronic equipment and holding to a minimum all course corrections. Also related to minimizing the number of course corrections is the shrinking supply of pressurized nitrogen for the craft's thrusters, whose proper operation is dependent on adequate gas pressure.

A successful fly-by will provide astronomers and scientists with information about the two least-known planets of our solar system. How extensive are the rings of Uranus, and do they contain embedded moons and moonlets like the rings of Saturn? What does Neptune look like, and does it have rings? (From Earth it appears only as a fuzzy, ill-defined disc.) If Voyager II succeeds on its final mission, it will produce information that will require a decade for scientists to interpret fully.

Skylab

When launched in 1973, **Skylab** was the world's largest satellite. It measured about 118 feet long and 22 feet wide (Figure 9.4). Larger and heavier than a typical three-bedroom house, it weighed 85 tons.

Actually, Skylab was a large and well-equipped laboratory capable of conducting extensive research in many fields of science. One of NASA's principal objectives in launching it was to prove that people could function effectively in outer space over long

Figure 9.4 Skylab The world's largest spaceship, Skylab, rapidly lost orbital speed and altitude due to increasing atmospheric friction. It fell to Earth in 1979. (Photo courtesy of NASA)

periods of time and perform assignments calling for mental concentration and physical stamina. At the same time, it tested a person's ability to live for long periods in confined quarters with two other people and to deal with the problems caused by weightlessness. No other space vehicle has ever offered astronauts such a wide spectrum of electronic equipment for conducting almost every kind of measurement and for carrying on nearly any type of research.

When Skylab was launched, NASA engineers forgot to take into consideration two critical factors in determining its life expectancy as a space vehicle. One was the onrush of electrical particles (electrons, protons, and other particles) continuously streaming from the sun. Scientists call this stream the *solar wind*, and it tended to push Skylab earthward from its 300-mile-high orbit. The second factor was the present trend toward a slight increase in radiant heat from the sun, which expanded the earth's atmosphere and caused a noticeable increase in the drag on Skylab. Scientists originally thought Skylab would last long enough to be rescued by a crew of astronauts and a booster rocket carried aloft by NASA's first space shuttle, Columbia. However, Columbia was still not operational, and Skylab continued its flight as its orbit dropped lower each week. Finally it lost critical speed, disintegrated over the Indian Ocean in July 1979, and plunged to Earth in Australia's sparsely settled outback. The crash caused no loss of human life and only minor property damage. Between its launch in 1973 and its disintegration and plunge to Earth in 1979, Skylab was staffed by astronauts for a total of only 171 days. By comparison, its Soviet counterpart, Salut VII, although far smaller and offering crew members a more austere environment, has been almost continuously manned by Russian astronauts since its launch.

Skylab's role as a space laboratory has been partially assumed by the operable space shuttles, although they offer less space for experiments and provide less technical equipment. However, on space shuttle mission 9, a smaller, European-built version of the original Skylab will be placed in orbit. Known as Spacelab, it will be operated by a multination consortium.

Pioneer X

In June 1983, after eleven years of flight, space probe Pioneer X crossed the orbit of Pluto, our most distant planet, and became the first man-made object to leave our solar system. To scientists and electronics engineers, the entry of Pioneer X into interstellar space was most exciting. Unlike the probes of the Voyager missions, Pioneer X's electronics were powered by a miniature atomic reactor rather than by solar panels. Theoretically, the spacecraft can travel for countless years into interstellar space and still maintain operation. The question confronting NASA engineers is how much longer reliable electronic communications can be maintained with Pioneer X and its 8 watt transmitter, already 2.8 billion miles from the sun. Engineers estimate that the satellite will remain in contact with its earth

station until it is 5 billion miles from the sun—a lonely and perhaps uneventful odyssey that will require at least a decade of constant travel.

Space Shuttle Orbiters

At first glance, the **space shuttle orbiter** resembles a commercial jet aircraft about the size of a Douglas DC-9 airliner, but with a thicker midsection (Figure 9.5). However, the shuttle has no conventional jet engines. Instead, mounted at the rear of the tail section are three rocket thrusters used to control the spaceship once it has been launched into orbit. The term *shuttle* derives from one of the craft's major objectives: shuttling aloft payloads of communications satellites, ready for service, and returning to Earth satellites in need of repair or readjustment. With a load capacity of 65,000 pounds, the shuttle can readily handle these assignments.

Budget Considerations

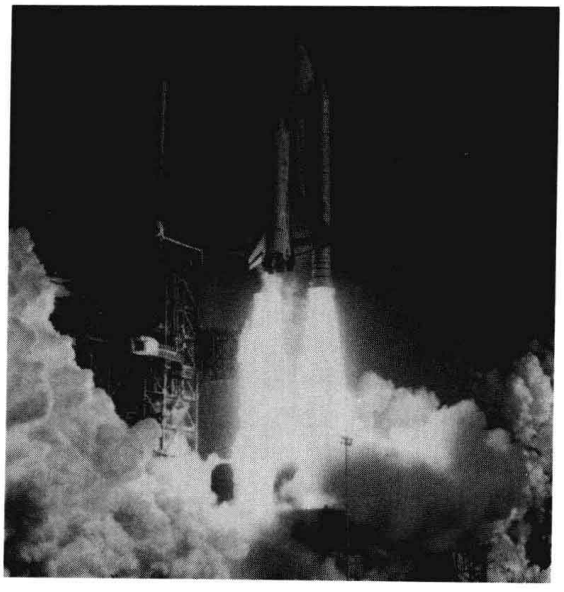
Important to NASA is the fact that the space shuttles are reusable. Each spacecraft can be used many times if properly serviced after each flight. Columbia has made five successful flights since its initial launch in April 1981. The major cost item in a space shuttle launch is the booster rockets and their fuel. NASA hopes to reduce this cost by devising better methods to salvage and recondition the fuel tanks. At present, they are parachuted to the sea when empty and then sold as scrap metal.

In NASA's 1983 budget request of $6.6 billion, nearly half was allocated to making the shuttle the "workhorse" of civil and military efforts in space into the 1990s. However, other traditional NASA activities, such as planetary space probes, astronomical studies, and pure research, will be continued at a somewhat reduced level.

Space Shuttles' Functions

Although the shuttles are much smaller than Skylab, they do provide scientists with elaborate electronic equipment for pure research, particularly

Figure 9.5 Space Shuttle Enterprise A sister ship, Columbia, was the first space shuttle launched into orbit on 12 April 1981. The shuttle is larger than many commercial airliners. (Photo courtesy of NASA)



Space Shuttle Columbia rises from launch pad at Cape Canaveral on 12 April 1981. Columbia has completed five missions and is scheduled for many more. (Photo courtesy of NASA)

under conditions of weightlessness. On board are scores of highly specialized sensors—essentially transducers that convert one kind of energy to another, usually into a varying electrical current or voltage. In digital form the data are continuously transmitted by telemetering to the mission's nearest ground station and stored in computer memory for later recall and study. Besides sensor studies of phenomena such as radiation and magnetism, the shuttles also have facilities for limited experimentation with small-scale production of sophisticated alloys and pure crystals in a weightless environment.

Although space shuttle astronauts can take over and "fly" the craft at any time, flight is normally controlled by four on-board computers overseen by a ground crew of four specialists at the space center. Usually two astronauts are on duty at the control position to monitor the ship's flight on screens similar to TV screens.

One of the major reasons for the development of the space shuttle series was the need for a vehicle that could both launch new satellites and retrieve inoperative satellites in need of repair. On 13 November 1982, on its fifth successful mission, the crew of Columbia announced that they had launched without incident a second communications satellite from its cargo hatch. The first satellite, launched the day before, was for Satellite Business Systems. The second was a communications satellite placed in space at the request of the Canadian government. Both were later positioned in geocentric orbit some 23,000 miles above the equator.

Future of Space Exploration

NASA's budget request for 1983 revealed modified but not completely changed objectives. The manned space flight to Mars and the radar-imaging space probe to Venus were dropped, but increased emphasis was placed on surveillance and communications satellite technology.

Space shuttles remain an integral part of NASA's plans. Because the thrust of NASA research and technical support programs is focused on the space shuttles, close adherence to their proposed launch schedule is important. The shuttles Columbia and Challenger are expected to fly further missions. Plans for Columbia include taking aloft **Spacelab** (Figure 9.6), a new and smaller version of Skylab.

NASA has a history of close cooperation with NOAA in the design, supervision of construction, and launch of NOAA's highly specialized satellites. Future launches for NOAA will employ one of the reusable space shuttles.

NOAA

Continuous surveillance and careful study of Earth's oceans, atmosphere, and land masses are the principal responsibility of the National Oceanic and Atmospheric Administration (NOAA), a division of the Department of Commerce. To accomplish its mis-

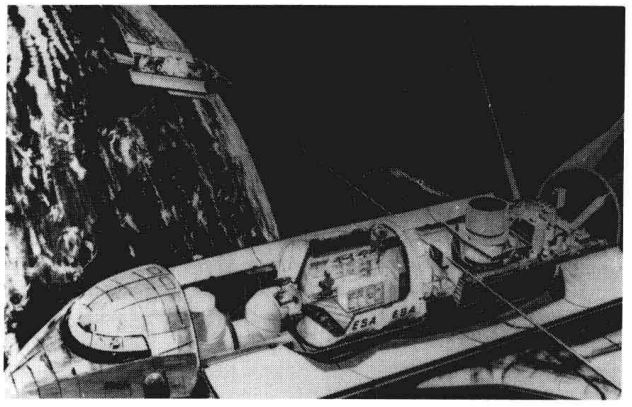


Figure 9.6 Spacelab To replace the original Skylab that fell to Earth in 1979, NASA proposes to launch the new, smaller Spacelab during mission 9 of the space shuttle. Photo shows an artist's concept of how the lab appears as the cargo bay of the shuttle opens. (Photo courtesy of NASA)

sion, NOAA employs highly sophisticated satellites and advanced electronic systems, most of which have been developed jointly with the staff of NASA. Originally it was envisioned that NOAA would grow to a size approaching that of NASA. Although that growth has not yet been realized, both NOAA and NASA are constantly working on joint projects.

The apparently similar roles of the two agencies may cause confusion. Both employ satellites in their data-gathering operations. Both use advanced electronics to command and control their satellites and more electronics systems to send data back to ground stations. At first glance, NASA and NOAA may seem to be performing overlapping functions. However, NASA is primarily research oriented, while NOAA is primarily concerned with operational systems. NASA does much research in earth-remote sensing systems. In collaboration with other agencies, it operates a laboratory for atmospheric sciences and conducts advanced research in specialized divisions such as earth survey applications and earth observation systems. NOAA operates these systems as they are developed, processes the data they provide, and evaluates and interprets the data for its own use and the use of other government agencies.

NESDIS

Within NOAA is a separate division accountable for the continuing operation of the agency's data-gathering satellites and their related electronic systems. The National Environmental Satellite Data and Information Service (NESDIS) is responsible for overseeing the total satellite service of NOAA, including regular checking of satellite orbits. NESDIS is also responsible for all electronic systems used for control, data gathering, and data transmission to ground stations. Currently, the two principal earth stations operated by NESDIS are at Wallops, Virginia, and Gilmore Creek, Alaska. Much data, largely in digital form, is constantly pouring into these stations and being relayed to the agency's central processing facilities at Suitland, Maryland, for evaluation.

Weather and Environmental Satellites

Weather satellites fall into two categories: **geosynchronous** (geostationary) and **polar orbiting.** Geosynchronous satellites are in synchronism, or in step, with the earth's rotational speed (about 1000 miles per hour at the equator) and consequently are always stationed over the same point on the earth.

The typical geosynchronous satellite (Figure 9.7) is at an altitude of a little less than 23,000 miles above the equator and has an orbital speed of about 7000 miles per hour. This situation represents the

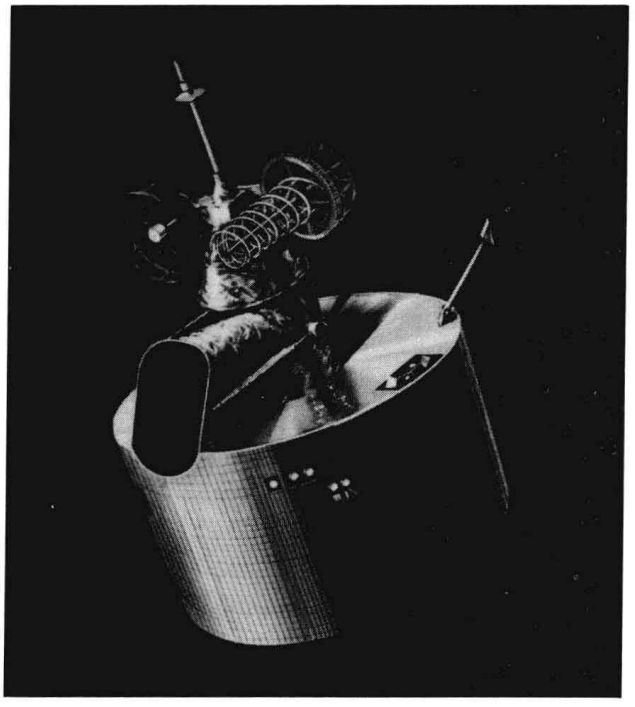


Figure 9.7 Geosynchronous Satellite Stationed nearly 23,000 miles above Earth's equator, the satellite's infrared scanner can readily cover an entire continent. Structures above the drum are mostly antennas. The drum is covered with silicon photovoltaic cells to power the electronics. (Photo courtesy of NOAA)

best compromise between long orbital life for the satellite and the maximum path loss that can be tolerated for radio transmission and reception. Currently, two NOAA geosynchronous satellites provide detailed weather conditions over the United States, Canada, Mexico, and much of South America. One is located at 75° west longitude (over Brazil), the other at 135° west longitude (over the eastern Pacific). Eventually, five or six geosynchronous satellites will provide worldwide coverage. At present, Japan, India, and a consortium of European nations also operate weather satellites.

Since geosynchronous satellites stationed over the equator are viewing the north and south polar regions at a very acute angle, no detail of any consequence is visible at latitudes 70° or higher. Therefore, NOAA uses a pair of rapidly moving satellites in low-altitude paths. They are called polar-orbiting satellites (Figure 9.8). At an altitude of just 500 miles, a pair of orbiters provide superbly detailed pictures of cloud cover and weather fronts, and the polar areas, where most weather systems develop, are thoroughly covered. Two polar-orbiting satellites in constantly shifting paths can scan almost every point on the earth's surface at least four times in a twenty-four-hour period.

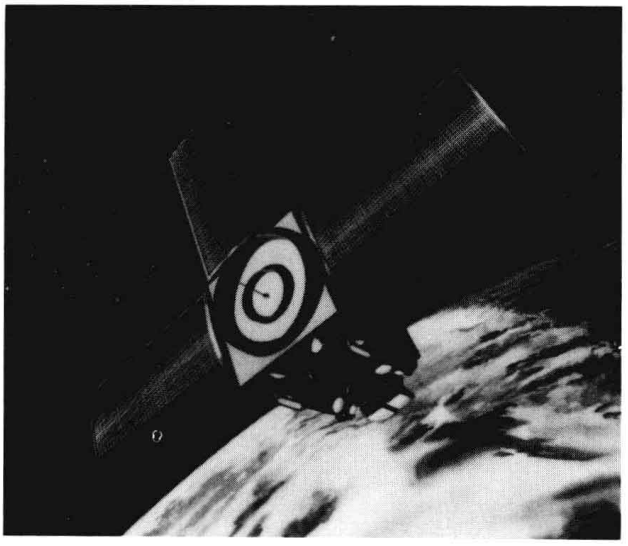
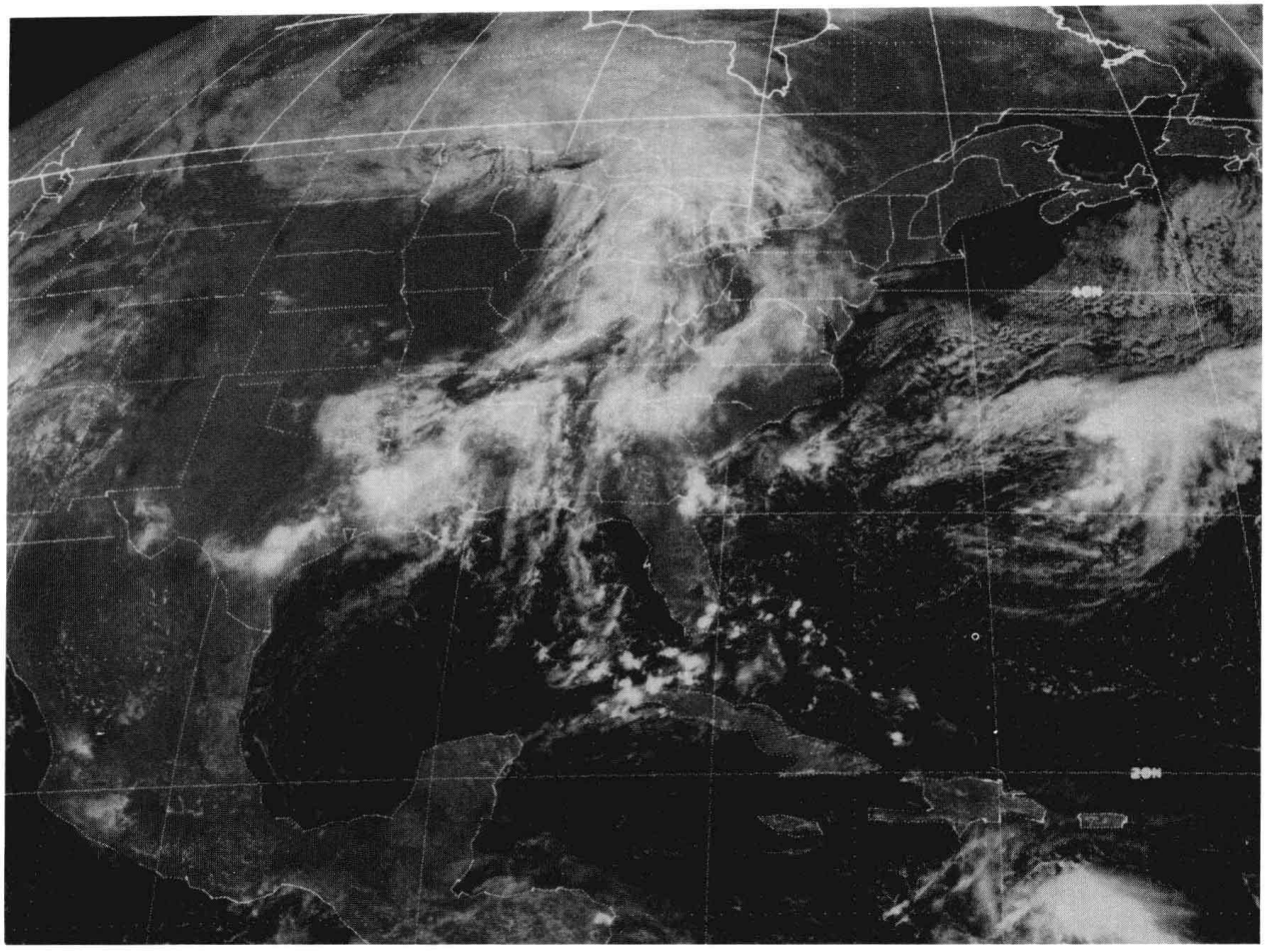


Figure 9.8 Polar-Orbiting Satellite Unlike the geosynchronous satellites that hover over the same position on the earth's surface, the polar-orbiting satellites move in a progressive orbit and at a high speed from an altitude less than 1000 miles. (Photo courtesy of NOAA)

The geosynchronous satellites used to obtain data on cloud cover, movement of weather fronts, and storm development over very large areas are named for the video-sensing unit they employ: the visible and infrared spin scan radiometer (VISSR). The polar-orbiting satellites are named for their sensing equipment unit, called an advanced very high resolution radiometer (AVHRR). These units provide the same basic type of picture as the previously used Vidicon TV cameras but yield much more finely detailed views. More important, they are capable of taking excellent images with high resolution in complete darkness since any warm object (clouds, water, land) gives off varying amounts of infrared rays. Data from the video units are supplemented by regularly radioed reports received by the satellites from scores of land- and sea-based fixed stations.

Although infrared light in photography can cut through haze and smog, clouds remain completely opaque. Polar-orbiting satellites experience this problem when the earth's surface is hidden by a cloud layer. They use microwave transmissions that punch through the cloud layers and are reflected back to the satellite from the earth and sea below. The time required for the return of a transmitted pulse is primarily determined by the distance traveled, but such factors as temperature and humidity also affect travel time.

The constant flow of weather data from the satellites is sent to central processing facilities at Suitland and Camp Springs, Maryland, where meteorologists prepare daily and long-range forecasts. These forecasts are distributed to the nation's wire services and the electronic media, as well as to NOAA's network of nearly 350 FM radio stations that continuously transmit local and regional weather forecasts and severe-weather warnings. Designed to



In this close weather view of the United States and Central America, taken at night with an infrared radiometer, a disturbance is moving out to sea from over New England. Another disturbance is forming in the western Gulf of Mexico. (Photo courtesy of NOAA)

serve a radius of about forty miles from the transmitter, NOAA weather stations operate on an assigned frequency of 162.40, 162.475, or 162.55 megahertz. They use comparatively low power and the narrow-band FM mode of transmission. The weather forecasts alone can be picked up with an inexpensive, battery-powered receiver. More expensive receivers come equipped with the severe-weather alarm.

Landsat and the Multispectral Scanner

On 31 January 1983, NOAA assumed management of Landsat D, the fourth of the NASA-launched low-orbit satellites. Landsat D is specifically assigned to land exploration, and the data obtained are oriented to the special needs of government agencies and privately owned companies. The data are available

as both computer-compatible tapes and photographic images for a nominal fee.

A key element of the Landsat program is its use of highly specialized electronic sensors. All are equipped with a multispectral scanner that acquires data simultaneously from four bands of the visible and near-infrared portions of the spectrum. By combining several bands on color film through use of filters, false-color composite photos are produced that provide much information not obtainable through ordinary photography. Such data have been helpful in evaluating many diverse phenomena: the vigor and growth of field crops, the impact of insect attacks or drought on forests, the presence of underground water, and the geologic clues that usually indicate the presence of mineral resources and petroleum not readily seen from the earth's surface. A thematic mapper currently under NASA's development for future NOAA satellites will provide even more highly detailed information about the earth and its waterways.

SUMMARY

Charged by Congress with responsibility for space exploration, NASA has relied heavily on both unmanned and manned spacecraft for gathering data. Because of the expense involved in sending manned space probes into deep interplanetary space, during the last two decades all space probes have been of the unmanned type. They obtain their electrical power from panels of photovoltaic cells. Their operations are controlled from the ground by electronic communications focused on the spaceship through a large parabolic antenna, or dish. The same antenna receives data, encoded in digital form, from the spaceship's cameras and sensors.

Since the early 1980s, NASA has emphasized development of space shuttles, reusable spacecraft capable of carrying a dozen crew members plus a great deal of electronic equipment for research. Unlike previous manned probes, space shuttles generate their own electrical power without the use of panels of solar cells. They always orbit the earth at a comparatively low altitude. NASA is developing plans for a manned space station to be operational within the next decade.

NASA technicians designed and operated all-electronic cameras that have produced remarkably detailed views of the planets and their moons. NASA's demands on the Electronics Industry have resulted in production of smaller, lighter, and more reliable equipment and components that have had a profound effect on the electronics marketplace.

The research program of NOAA is largely directed to intensive studies of the earth's atmosphere, seas, and oceans. The earth's land masses are also included increasingly in NOAA's studies. NOAA daily acquires large amounts of data from its highly sophisticated satellites and from sensor-equipped radio stations on land and sea that report findings automatically. Weather data is obtained and processed from geosynchronous satellites some 23,000 miles above the equator and from rapidly moving polar satellites in orbits just a few hundred miles high. The data are used in preparing daily and long-range weather forecasts issued by the National Weather Service. Other satellites in low orbits provide information on forest cover, field crops, land erosion, and other natural phenomena.

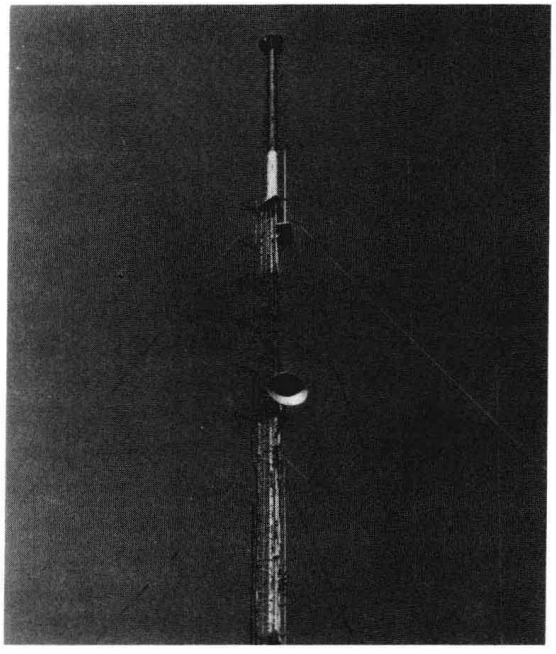
KEY WORDS AND CONCEPTS

geosynchronous satellite
parabolic reflector
polar-orbiting satellite
Skylab
Spacelab
space probe
space shuttle orbiter

SELF-TEST QUESTIONS

1. What is the source of electrical power used to operate the electronic equipment of all space vehicles?
2. What additional electronic equipment is always provided if the spacecraft is manned?
3. Why is it essential that spacecraft be equipped with highly directional transmitting antennas?
4. NASA's Goldstone ground station uses a parabolic antenna over 200 feet in diameter for reception and transmission. Why is such a large antenna required?
5. Describe the impact of NASA's space program on the quality of equipment and components manufactured by the Electronics Industry.
6. Since all space probes are unmanned, how are scientific data obtained for transmission to the NASA ground stations?
7. What two factors were largely responsible for Skylab's premature plunge to Earth?
8. What observations and measurements could astronauts make from Skylab's 200-mile-high orbit that were not possible or practical from the surface of the earth?
9. What two objectives did NASA scientists hope the space shuttle would accomplish?
10. How are NASA facilities and personnel used to aid private firms engaged in telecommunications?
11. Explain what is meant by a geosynchronous satellite.
12. Why do satellites equipped with infrared radiometers provide superior images of world weather and cloud cover?
13. Why does NOAA still use orbital satellites traveling at high speeds in low orbits to obtain data for weather forecasts?
14. What specialized information is best obtained from satellites equipped with microwave sensors?

10 Electronics in Tele-communications



In this installation in Eugene, Oregon, a single tower serves many applications. Shown are the TV antenna of KEZI-TV (top), the antenna of KPNW-FM (below left), and the microwave dish of KEZI-TV that relays programs from a studio on the valley floor to the transmitter site. (Photo courtesy of KPNW-FM; tower property of station KEZI-TV)

Historical Perspective

AM Broadcast Stations
- Classification of AM Stations
- Advantages of AM Transmission
- Job Assignments in AM Stations

FM Broadcast Stations
- Development of FM Transmission
- Growth of FM Broadcasting
- Classification of FM Stations

Television Broadcasting
- Early Television Technology
- How Television Works
- Emergence of TV Standards
- Color Television
- Television Job Assignments

Cable Television
- Use of Geosynchronous Satellites
- Cable TV Installation
- Rooftop Reception of Satellite TV

Future of Radio and TV Broadcasting
- Future of AM Stereo
- Future of Commercial TV Stations

Amateur Radio
Contributions of the Amateur Radio Service
Amateur Radio Today
OSCAR Space Satellites
American Radio Relay League

Citizens Radio Service
Problems Facing CB Radio
Radio Emergency Associated Citizens Teams
CB as a Training Ground for Amateur Operators

Summary

Key Words and Concepts

Self-Test Questions

Telecommunications means communication over a distance. Since the focus of this book is on electronics rather than on electrical communication, we confine our discussion to areas where telecommunication is accomplished by the generation, transmission, interception, and detection of electromagnetic waves. (Electrical communication makes use of metallic conductors rather than electromagnetic fields.)

Telecommunications fields discussed in this chapter include the Broadcast Services (AM, FM, and TV broadcasting) and the Personal Services (amateur radio and citizens band). Also administered by the FCC are the Safety and Special Radio Services, which include such areas as aviation, marine, public safety, and industrial and land transportation. Their estimated station count is near 700,000. The FCC also administers the Common Carrier category, including microwave, international telephone and telegraph, and satellites, with a station count of somewhat over 20,000. Discussion of these specialized services is beyond the scope of this text.

HISTORICAL PERSPECTIVE

Radio communication got its start in the United States when it was used in saving lives at sea. Thus, the Radio Act of 1912, the first definitive piece of legislation in the field, dealt in detail with the allocation of frequencies between government services and private firms so as to minimize interference with distress signals from ships on the high seas. The act also established a system of licensing radio stations and radio operators. The system was to be administered by the secretary of commerce and labor.

Virtually all radio communication in the early 1900s was in the form of radiotelegraphy, or **wireless.** Transmitters were usually of the high-voltage, electric arc type, and the receiving equipment was seldom more complex than a simple crystal set. De Forest had invented the triode radio tube in 1907, but the manufacture of vacuum tubes and other receiver and transmitter components on a sizable scale did not get under way until the communications emergencies created during World War I.

When Herbert Hoover was appointed secretary of commerce in 1920, there were only 2 broadcast stations on the air and about 40,000 receivers in the homes of listeners. A year later there were 300 stations and 2 million receiving sets. By 1922 nearly 600 broadcasting stations were in operation, their expansion stimulated by the enormous growth in the number of listeners. The public's appetite for entertainment via the airwaves seemed insatiable. Radio dealers were unable to keep up with the demand for receivers, and new radio stations were making their debut almost daily.

By 1926 the situation had grown intolerable. Listeners were frequently unable to receive even a single program without constant interference from other stations operating on the same or nearby frequencies. Hoover lacked the authority under the Radio Act of 1912 to assign stations to new frequencies and powers of operation that would minimize interference problems. At his insistence Congress drafted legislation known as the Radio Act of 1927, which created a Federal Radio Commission with authority

over all forms of radio communications, including radio stations and operators. The legislation reflected Hoover's view that the radio spectrum belonged to the people, not to the station owners. The act stated that the people have the right to lease frequencies for short periods of time to responsible individuals or organizations in the form of temporary assignment of station licenses if such licenses are in the "public interest, convenience, and necessity."

By the early 1930s Congress realized the necessity of regulating all electrical and electronic communications media, not just radio. The Communications Act of 1934 created the seven-member Federal Communications Commission, which has authority to regulate all aspects of diverse telecommunications segments: radio and television stations, cable TV systems, amateur and CB radio, telephone and telegraph wire communications, and satellite communications. All services discussed in this chapter are regulated by the FCC. In 1982 and 1983, the Communications Act of 1934 and its derivative rules and regulations were reviewed by the FCC to comply with a congressional request for deregulation of the agency's tight control over telecommunications.

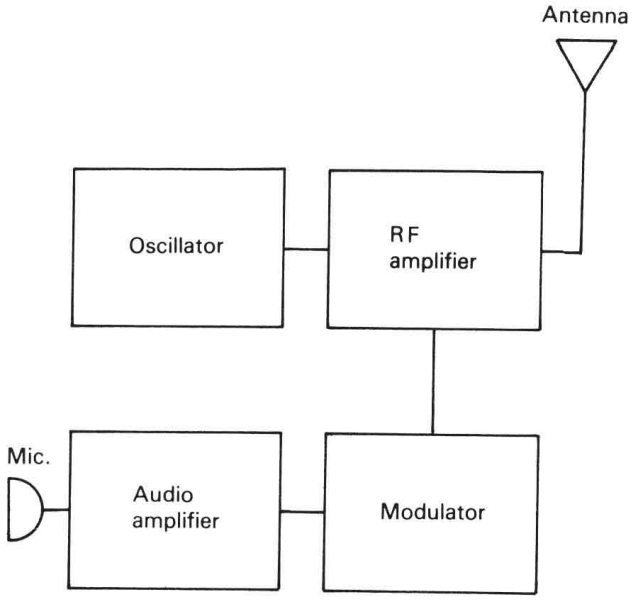


Figure 10.1 Block Diagram of a Typical AM Transmitter The oscillator generates the radio frequency wave; the RF amplifier increases its power. The audio amplifier increases the microphone's signal; the modulator further amplifies the signal and varies the power output of the RF amplifier at a voice or music frequency rate.

AM BROADCAST STATIONS

Originally limited to only two channels, 830 kilohertz and 750 kilohertz, the standard broadcast band was expanded in 1924 by Secretary of Commerce Hoover to include all frequencies between 1500 and 550 kilohertz. Later the high-frequency end was extended to include those frequencies between 1500 and 1550 kilohertz. At the conclusion of World War II, the high-frequency limit was extended to 1600 kilohertz, and one channel was added at the low-frequency end of the band to extend it to 540 kilohertz.

Radio stations operating within the 1600 to 540 kilohertz allocation are said to be in the *standard broadcast band,* and under terms of their FCC license, all must use **amplitude modulation** (AM). Amplitude modulation means that the power of the radio frequency wave generated by the transmitter varies at a voice or music frequency rate, while the frequency of the wave remains constant. Figure 10.1 shows a block diagram of a typical AM transmitter.

Classification of AM Stations

Since 1939 the FCC has classified all AM stations into four categories with subclassifications in each. The elite are the Class I stations operating on the so-called clear channels, usually with 50,000 watts of power (the maximum power allocated to any U.S. station). No other station in the United States is permitted to operate on their frequency at night. Class II stations operate with a maximum power of 50 kilowatts (although the minimum can be as low

as 10 kilowatts) as secondary stations on a clear-channel allocation. At night they must reduce power, use a directional antenna to minimize interference to the dominant station assigned to the channel, or in some cases leave the air. Class III stations are known as regional stations since they serve a considerable geographical area, not just a town or city. They usually run at least 5000 watts daytime power and reduce their signal interference to the dominant station on their channel at night by reducing power to 1000 watts, by using a directional antenna to steer the signal in some other direction, or by doing both. Class IV stations serve purely local areas such as towns or small cities. They operate with a maximum daytime power of 1000 watts and a minimum nighttime power of 250 watts. They operate without a directional antenna. In your community you may find AM stations operating in two or three of these categories.

A recent edition of the *Broadcasting Yearbook* lists 4559 AM stations operating in the standard broadcast band. Since interference between stations is already serious, it is unlikely that the FCC will permit the number of AM stations to grow beyond this figure.

Advantages of AM Transmission

Probably the greatest advantage of AM transmission is its simplicity of operation. No critical adjustments need to be made, and a minimum of technical equipment is required. In the past, AM has been widely used by amateur stations and stations engaged in the safety and special radio services. Currently AM is used by nearly all CB operators, and TV stations use an AM transmitter to broadcast the visual information and the controlling pulses.

On the other hand, the audio quality obtainable from an AM broadcast station is appreciably inferior to that of an FM broadcast station, although the difference may not be apparent to the casual listener. The 10 kilohertz channel width restricts the highest frequency audio tone that can be transmitted to about 10,000 hertz. Also, the standard broadcast band is extremely crowded—so crowded that in recent years the FCC has been returning all applications for new AM stations and suggesting that applicants refile for FM facilities. Crowding makes it nearly impossible to receive a station's programming without interference from other AM stations on the same or nearby frequencies. Static discharges, lightning flashes, and electrical equipment that is arcing or sparking also create radio waves with an AM characteristic, which are naturally picked up on an AM receiver tuned to an AM station.

However, a well-adjusted AM transmitter using high-quality program material is capable of excellent audio quality, as evidenced by the sound heard on the control-room loudspeaker. The weak link in the chain is the average AM broadcast receiver with its poorly designed audio system and its totally inadequate loudspeaker. Many of the cheaper sets are hard pressed to deliver a frequency response from 500 to 3500 hertz. Thus, they may sound tinny and unnatural.

Job Assignments in AM Stations

In a commercial radio station the principal job categories are sales, programming, and engineering. The sales staff provides the station's income. In a well-run, community-oriented station, about 75% of the station's gross income will be derived from the sale of spot announcements to local merchants. The balance will come from regional and national advertising. The programming department is responsible for everything the station broadcasts: music, news, sports, public-service announcements, and the like. (In the larger stations news may be a separate department.) Engineering is headed by a chief engineer, a person trained in the installation and maintenance of all types of electronic equipment.

FCC rules adopted in 1980 state that a chief engineer must hold the FCC General-Class Radiotelephone Operator's license. (Formerly, the Radiotelephone First Class was required.) The Gen-

In a typical AM broadcast station, a newscaster uses prerecorded tapes of network news and local interviews plus teletype wire service copy in presenting scheduled newscast. (Photo courtesy of KPNW, Eugene, Oregon)

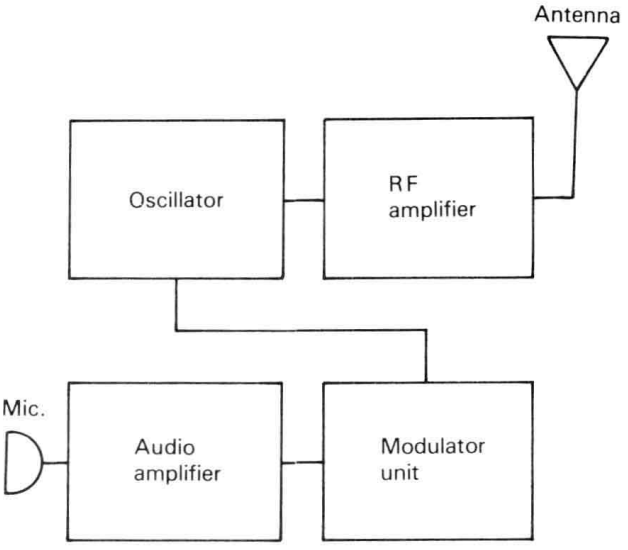


Figure 10.2 Block Diagram of a Typical FM Transmitter The oscillator generates the radio frequency wave; the RF amplifier increases the wave's power. The audio amplifier increases the microphone's signal; the modulator varies the oscillator frequency at a voice or music rate.

eral-Class license is approximately the same as the older Radiotelephone Second-Class certification. The chief engineer and the station's licensee are responsible for station operation in compliance with the rules and regulations of the FCC and the terms of the station's license. Job assignments and responsibilities on FM stations are similar to those on AM stations.

FM BROADCAST STATIONS

Frequency modulation (FM) was the brainchild of Major Edwin Armstrong. During World War I he also invented the superheterodyne receiver circuit, which is incorporated into the design of all AM, FM, and TV receivers. Armstrong was engaged in a prolonged legal battle with Lee De Forest over the priority of his patent, which described the use of De Forest's audion, or triode vacuum tube, in an oscillatory circuit. Armstrong eventually lost the decision after years of bitter litigation. Today he is principally remembered for his development of frequency modulation, a new system of broadcasting and communication. A frequency-modulated station is one in which the power output remains constant at all times, while the carrier wave's frequency varies at a voice or music rate. Figure 10.2 shows a block diagram of a typical FM transmitter.

Development of FM Transmission

Armstrong's original approach to frequency modulation was based on his effort to find a noise-immune system of broadcasting and communication. He noted that such common forms of electrical interference as static discharges, lightning flashes, automotive ignition, electric motors with sparking brushes, and leaky insulators on power lines all produced radio waves with a varying amplitude (AM) characteristic. He resolved to produce a system almost completely

insensitive to amplitude changes and extremely sensitive to frequency variations (Figure 10.3). He started his research by working with his former instructor in electricity at Columbia University, Dr. Michael Pupin. Later he elected to set up his own laboratory for FM research and development.

Engineers were already familiar with the principles of frequency modulation. Early attempts to amplitude-modulate an AM transmitter designed for the standard broadcast band frequently produced a carrier wave with both amplitude and frequency variations. When received on a conventional broadcast receiver, voice and music sounded badly distorted because of the frequency variations.

Through experimentation Armstrong concluded that the maximum frequency swing of the carrier wave on each side of the assigned frequency was to be 75 kilohertz, which was to correspond to maximum power transmission (or 100% modulation) in a standard broadcast transmitter. This variation, or deviation, is usually written as ±75 kilohertz and requires that the FM channel be 200 kilohertz wide to prevent interference with adjacent channels. Around this basic standard Armstrong developed both transmitting and receiving equipment. The original allocation for FM broadcasting was 42 to 50 megahertz, which proved inadequate. It was later moved to the radio spectrum lying between VHF TV Channels 6 and 7, 88 to 108 megahertz, where it remains today. Channels are numbered from 200 to 296, each 200

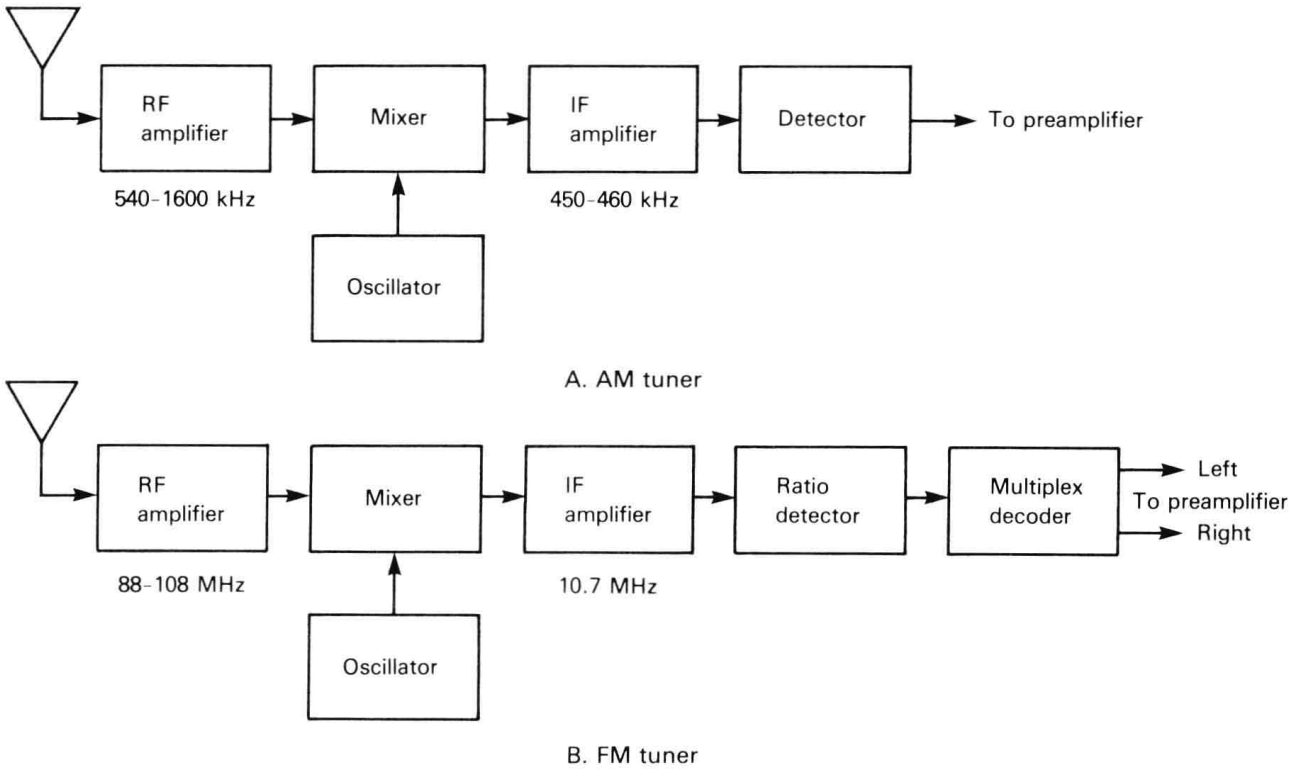


Figure 10.3 Block Diagram of an AM/FM Tuner Unit The detector in the AM tuner responds only to amplitude modulation, while the ratio detector in the FM tuner responds only to frequency variations.

kilohertz wide, and channels 200 to 220 are reserved for educational FM stations.

Growth of FM Broadcasting

Although Armstrong demonstrated his new FM system before a panel of engineers as early as 1935, the timing was unfortunate. RCA and Dumont were more interested in developing television. RCA, particularly, viewed FM as a threat not only to its tremendous investment in the manufacture of AM transmitting and receiving equipment, but also to its ownership of AM radio stations. Even when World War II was concluded and FM transmitters and receivers were once more being produced and distributed, only a small number of receivers were in the hands of listeners.

Since the 1960s, FM broadcasting has grown more popular, stimulated by its adaptability to high-fidelity stereo transmission and reception. A secondary factor is the development of specialty program formats such as sports, jazz, and classical music, which meet special listener interests not served by conventional AM broadcast stations. By the mid-1970s in most communities of 100,000 people or more, every locally available channel was in use or had already been applied for.

Classification of FM Stations

FM stations are classified into three divisions based on the area the station serves and the power of operation. Class A stations are limited to 3 kilowatts of power and serve a relatively small community or area. Class B stations are assigned to sizable communities and may use up to 50 kilowatts of power. Class C stations serve both a community and a large surrounding area and may use power up to 100 kilowatts of power. A recent edition of *Broadcasting Yearbook* lists 3155 commercial and 1038 educational FM stations currently in operation.

TELEVISION BROADCASTING

Early Television Technology

Although widespread use of television as an entertainment medium dates back only to the early 1950s, the first attempts to transmit pictures over a wire or cable took place in 1875, when C.R. Carey of Boston developed the first practical system. In principle, his system was similar to that employed in the animated electrical signs currently found in sports stadiums or the "traveling signs" used for news flashes and advertisements.

Carey's receiver was a panel made up of many small, low-voltage lamps. One terminal of each lamp was brought out through a separate lead to an electrical relay controlling the application of battery power. The transmitter was made up of a corresponding panel of photoelectric cells. Separate leads from each cell were brought out to the relay windings.

When an image was projected on the bank of photoelectric cells, some were activated, and their relays turned on their corresponding lamps in the receiver unit. Cells that received insufficient light or no light failed to activate their relays. As a consequence, their associated lamps in the receiver did not light. The result was a shadow picture on the light panel—a rather crude reproduction of the original image projected on the photoelectric cells.

The mechanical delay in the operation of the relays and the absence of what photographers refer to as *gray range*—anything between black and white—made this approach impractical for the transmission of detailed pictures that created the illusion of motion.

Edison's Contribution

Thomas Edison's experimentation with sequential picture presentation—first, through the use of opaque flip cards and later, sprocketed photographic film to create the impression of pictures that moved—

helped define the problems television needed to surmount. At the same time his experiments pointed to some possible solutions.

Edison's experiments underlined several important characteristics of the human eye that are basic to any approach to sequential picture presentation. The most important of these characteristics is *retentivity of vision,* which means that a vivid image impressed on the eye's retina tends to persist for perhaps one-tenth of a second before dying out. If, before it completely fades away, it is replaced by a second vivid image, the eye can be tricked into believing that no interruption has taken place. If slight changes in the picture elements are made from one picture to the next, the illusion of motion can be created. Another important fact demonstrated by Edison's experiments is that a great deal of detail in the image is not necessary—the eye tends to supply the missing detail.

Nipkow's Contribution

Starting with basic information supplied by the earlier experimenters, the German scientist Paul Nipkow in 1884 decided upon an entirely different approach. His approach was to be the one we use today: Dissect the picture into bits and pieces (picture elements), transmit it, then reassemble it in the same sequence. If the pictures could be dissected and reassembled rapidly enough, the illusion of motion could be created. Nipkow's basic problem was the state-of-the-art electronics. Instead of a sensitive studio camera, he was forced to use a slow-acting and insensitive selenium photoelectric cell. In place of a picture tube, he had to be content with the varying light intensity of a neon tube. Although Nipkow used the technique of horizontal line-after-line scanning as used today, it required the use of whirling metal discs, with a precisely drilled helical pattern of holes, driven by synchronous motors. Nipkow's technique was the mechanical counterpart of today's all-electronic scanning system, but it was the right approach to television transmission and reception.

Zworykin's Contribution

The Russian-born scientist Vladimir Zworykin developed Nipkow's line-by-line scanning to a highly perfected electronic technology. As director of research at RCA, which was interested in developing an all-electronic system of television transmission and reception, Zworykin directed the expenditure of millions of dollars on a number of research and development projects. In 1938 he perfected the first television camera capable of quality picture reproduction. He called it an **iconoscope** (literally, an "image viewer").

How Television Works

Although modern camera tubes, known by such names as vidicons, plumbicons, and satyricons, are greatly improved in terms of sensitivity and ability to render fine detail, their basic principle is essentially the same as that of the iconoscope. A lens system focuses the image on a rectangular sheet of mica, one side of which is covered with millions of separated silver globules sensitized by exposure to the vapor of the rare metal caesium. This sensitized rectangle is known as the **mosaic.** An electronic scanning system provides a very narrow beam of electrons that scans the sensitized surface from left to right as it gradually works down from the top edge of the mosaic to the bottom. The scanning sequence is then rapidly repeated. A simplified block diagram of a basic television system is shown in Figure 10.4.

When the studio scene is brought into sharp focus on the mosaic, light, medium, and dark gray areas are created. The light areas have the most intense electrical charge; the medium gray areas a lower value of charge; the blacker areas, no appreciable charge. The scanning electron beam releases these charges in sequential order to produce a pulsating DC current in the external camera tube circuit. For the scan of a single horizontal line, the DC pulsations are an electrical analog of the variations in light intensity encountered along that same line (black through intermediate gray through white).

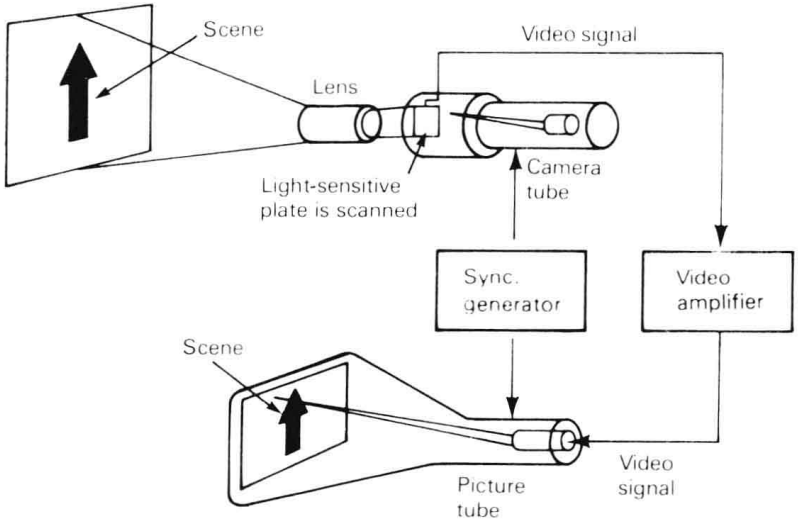


Figure 10.4 Block Diagram of a Basic Television System The studio scene is sharply focused on the sensitized mosaic (plate) of camera tube. An electron beam in the picture tube of the TV receiver scans its phosphor-coated screen in unison with the camera scan to recreate the camera image.

The camera signal, as the pulsating DC current is called, after amplification by a video amplifier circuit, can then be applied to the control grid of the cathode-ray tube used in a TV monitor. With both TV camera and TV monitor tubes supplied additionally with precisely timed controlling pulses that keep their scanning beams in exact synchrony, the scanning beam within the monitor tube systematically traces and retraces the picture imposed on the camera's mosaic. When the pulsating DC current pulse is strong on the cathode-ray tube, its phosphor coating glows an intense white; when weak, its phosphors give off some variant of gray; when it momentarily falls to zero, no phosphorescence occurs, and the tracing is invisible (black).

The net effect of this synchronized electronic scanning process between camera tube and picture tube is an exact reproduction in **monochrome** (black and white) of the original studio scene. The quality is somewhat less than perfect, but the human eye is easily tricked and tends to supply any missing detail. The original goal was to achieve a quality of reproduction equal to 16 millimeter motion picture film. This goal has been closely approached but never quite achieved.

Emergence of TV Standards

Largely through trial and error, standards for all-electronic TV were developed and given the status of law by the FCC. Borrowing from 16 millimeter motion picture technology, which eliminated the annoying flicker by projecting 24 separate pictures per second and projecting each two or three times, the U.S. TV system transmits 30 complete pictures per second—that is, **frames**—and 60 partial pictures, or **fields.** From the top of the picture to the bottom are 525 potential scanning lines. The odd-numbered lines are scanned first to provide one field. The even-numbered lines are scanned to provide the next field. The two combined provide one complete picture, or frame. The odd-even line-scanning technique reduces flicker to a very low value and is called **interlaced scan.** The U.S. standards prevail also in Canada, Mexico, and Japan. A 625 line standard

that is widespread in Europe greatly improves picture detail, particularly in color.

In the U.S. system, the picture information and the controlling pulses are transmitted by a very powerful AM transmitter. (The controlling pulses turn off the electron beam at the end of each line and at the end of each scanning field and maintain synchronism between the scanning lines of studio camera and home TV set.) A separate FM transmitter, frequently of considerably lower power rating, transmits the sound information. The outputs of the two transmitters are combined in a device known as a **diplexer** and then fed to a single transmitting antenna of the hertzian type (Figure 10.5). In Western Europe the sound is transmitted by an AM rather than an FM transmitter.

Because of the amount of information that must be sent by the transmitter that carries the picture signal (with its wide spectrum of electrical variations as it goes from black to white through countless shades of gray), TV channels need a channel assignment 6.0 megahertz wide. Most of this allocation is used for the picture and the DC pulse information. The sound channel requires only 50 kilohertz of spectrum to transmit all audio information in the frequency range 50 to 15,000 hertz as required by the FCC. One TV channel with its 6.0 megahertz allocation uses up as much spectrum as 30 FM stations or 600 standard AM broadcast stations. At present 746 commercial and 267 noncommercial TV stations are on the air in the United States.

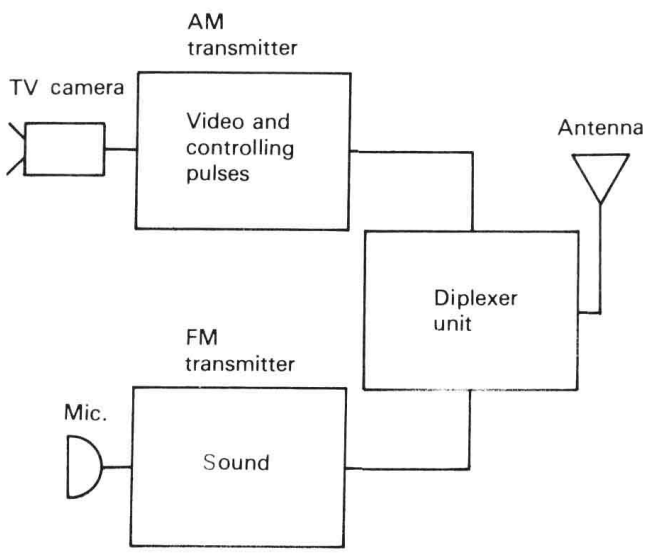


Figure 10.5 Block Diagram of a Television Transmitter Picture information is transmitted by an AM transmitter, sound information by an FM transmitter.

The typical TV receiver is a hybrid in design and construction. A tuner, common to both AM and FM sections, selects the desired channel and converts it to a new and lower frequency, the **intermediate frequency** (IF). After IF amplification, the picture and pulse (AM) signals are separately amplified and channeled to the video amplifier and control circuits. The sound signal (FM) is also separately amplified and channeled to the loudspeaker, which is usually too small to provide truly satisfactory bass response.

Color Television

The transmission and reception of color by television is far more complex than the transmission of black-and-white programs, and its technology is well beyond the scope of this text. However, the basic principle upon which the color system operates is as follows: By combining the three primary colors (red, blue, and green) in varying proportions, any other color can be produced. An equal amount of all three colors theoretically produces white.

At the TV station, the color camera is equipped with three camera tubes rather than with the single one used in black-and-white transmission. By the use of appropriate optical filters, one camera tube responds only to the red picture information in the studio scene; the second picks up and supplies blue picture information; the third provides green picture data. By complex means, the electrical signal from each camera is, in effect, separately transmitted to the color TV receiver. In the color set, the picture tube, instead of being a single cathode-ray tube as

in a monochrome receiver, is actually three cathode-ray tubes, each with a separate electron gun but mounted in the same glass enclosure.

Instead of a continuous phosphor coating on the inside of the face of the picture tube (as is the case in a monochrome set), the color picture tube's phosphor coating is made up of an extremely large number of tiny dots of red, blue, and green phosphor material deposited in a triangular pattern and positioned along horizontal lines with great precision. The three very sharply focused electron beams from the three electron guns, each varying with its corresponding camera signal, are carefully aligned, or converged, so that they only strike the correct set of phosphor dots. A perforated metal screen is positioned near the phosphor coating to insure that stray electrons from the red electron gun, for example, will not accidentally strike the blue or green phosphor dots and cause incorrect color rendering. The screen is called a *shadow mask*. Future tube models may dispense with it. Provided that the electron guns (red, blue, and green) are precisely aligned and adjusted in intensity, that the receiver has adequate bandwidth (6.0 megahertz), and that the fine-tuning control is properly set, an excellent reproduction of the studio scene should be obtained.

From the television director's booth the signal arrives at the master control position, where a technician keeps watch on the waveform of the camera signals and adjusts color balance. From master control, video and audio signals go directly to the visual and aural transmitters. (Photo courtesy of KVAL-TV, Eugene, Oregon)

Television Job Assignments

A television station usually employs four or five times the number of staff members as an AM or FM radio station. In a metropolitan area of moderate size, a TV station might have a staff of about fifty, while a radio station might have only ten to twelve.

In general, most broadcasting employees must complete at least two years of study in an accredited community college or university in some field that relates closely to their employment. For example, news staff members would be expected to have some college-level work in writing and journalism. Advertising and sales staff members should have studied public relations and the writing of sales copy. Engineering staff members must hold the new FCC General-Class Radiotelephone Operator's license and have advanced training in electronics and physics. Department heads should have earned at least a baccalaureate degree in business administration. Practical experience is more important to the prospective employer than formal training alone, but a combination of the two is most desirable.

In AM/FM broadcasting, staff members must be able to read advertising copy so that it sounds conversational. They also must be adept at writing advertising copy that sells. Further requirements are imposed on a TV station's program and news staff. Staff members must be reasonably photogenic and well groomed, be able to ad-lib when the need arises, and be able to read smoothly from the teleprompter just beyond camera range.

CABLE TELEVISION

Formerly called Community Antenna Television Systems (CATV), cable television now provides TV programs to about 15% of the 120 million TV sets in U.S. homes. Since cable services first started in the early 1950s, their basic objectives have been

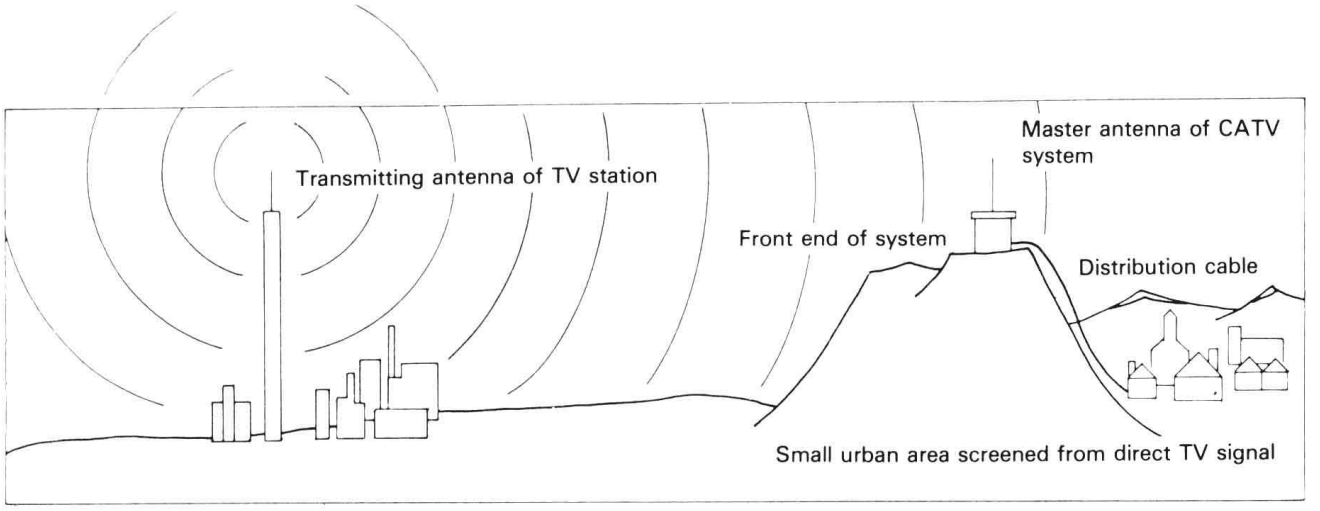


Figure 10.6 Cable TV System A community antenna television system (CATV) brings reception to an urban area screened from the direct signal by a high hill. Today cable TV systems rely heavily on direct pickup of satellite-relayed programs.

considerably modified. In the days when local TV stations were very scarce, particularly in the smaller towns, antennas were often set up on high ridges and mountaintops, where signals from stations located in the larger cities were well received (Figure 10.6). The signals were then amplified and distributed to TV set owners throughout the community over coaxial (shielded) cable. Subscribers paid a modest installation charge and monthly rental fee for this service. Figure 10.7 shows a block diagram of a typical cable TV pickup and distribution system.

By the 1960s new TV stations were being constructed, even in communities with populations of 50,000 or less, and some cable subscribers began to cancel their subscriptions and go back to using rooftop antennas. Cable companies wanted to hold on to their subscribers and at the same time attract new viewers. The first approach used in many areas was development of public access channels programmed from a local studio with technical equipment supplied by the cable company. Such programming was assigned an unused channel on the dial. In many communities locally produced programs dealing with public issues did stimulate TV watching, but a more dramatic change in program offerings was necessary if the cable companies were to expand and thrive, particularly in communities where several commercial TV stations were affiliated with major networks. By the late 1970s the technology had been developed to relay a whole spectrum of new programs, both live and recorded, by means of communications satellites owned by such firms as Western Union, AT&T, and RCA Americom.

Use of Geosynchronous Satellites

As we learned in Chapter 9, geosynchronous satellites are stationed in orbits above the equator at an elevation of a little less than 23,000 miles (Figure 10.8). From this altitude their antennas can provide an excellent signal to an enormous area (referred to as the satellite's footprint) that often includes much of the forty-eight contiguous states.

Geosynchronous satellites are equipped for program relay with a number of **transponders** that per-

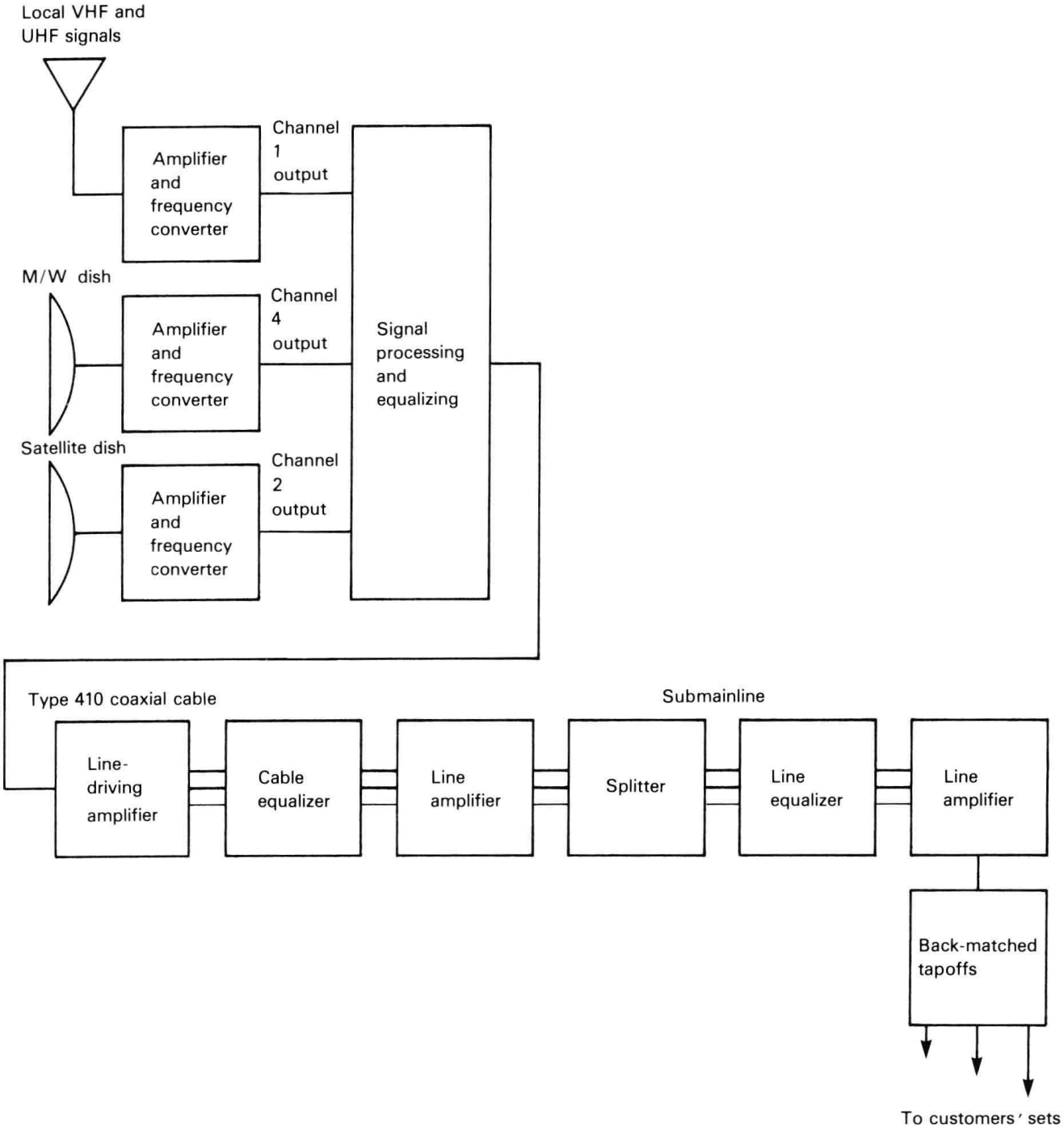


Figure 10.7 Cable TV Pickup and Distribution System The coaxial distribution cable is usually supported on public utility poles. The system usually requires a city or county franchise.

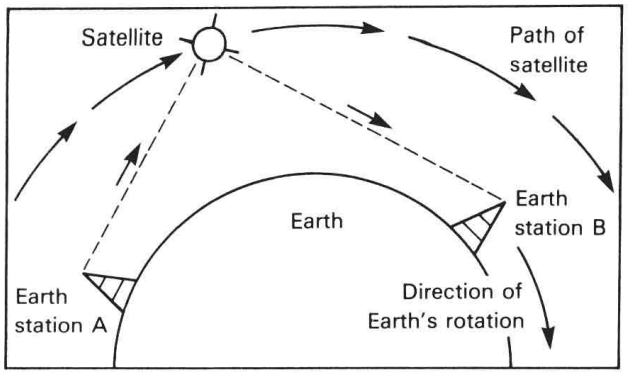


Figure 10.8 Geosynchronous Communications Satellite The satellite is positioned about 23,000 miles above the equator, and its speed is exactly paced to that of the earth (about 7000 mph compared to the earth's 1000 mph).

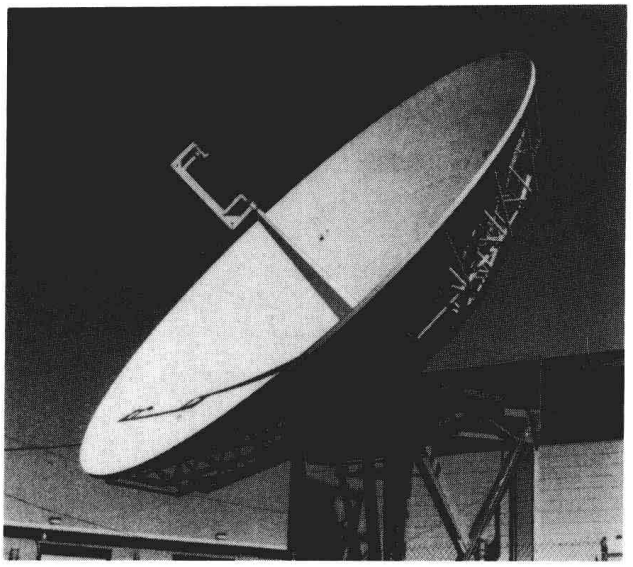


Figure 10.9 High-Gain Parabolic Receiving Antenna This 30 foot parabolic dish picks up specially produced TV programs relayed by a series of geosynchronous satellites. Aimed nearly due south, the dish is tilted at a vertical angle of about 40°. This antenna is used by Group W Westinghouse to provide programs to cable TV subscribers in Eugene and Springfield, Oregon.

mit the satellites to receive a signal from a ground originating station on one frequency and retransmit it on another. Use of transponders avoids the electronic instability that would result if satellite reception and retransmission of signals were to take place on the same frequency. The retransmitted signal is picked up by a high-gain parabolic receiving antenna, usually located on the premises of the cable TV company (Figure 10.9). There it is changed from a UHF to a VHF frequency to reduce losses, electronically processed, and sent out over the cable distribution amplifiers and cables as if it were a signal from a TV station.

Cable TV Installation

The cable technician first installs a converter unit, which works on the superheterodyne principle and changes all TV signals coming down the cable service drop to an unused channel—2, 3, or possibly 4—on an existing TV set. A remote control unit with twenty-five or more push buttons enables the viewer to select any channel by pushing the appropriate key, which changes the DC voltage applied to varactor diodes that serve as tuning capacitors. The set's tuning dial is left permanently set to the unused channel.

Premium channels, such as Home Box Office and Showtime, are visually scrambled, usually by removal of the sync pulses. For a flawless picture on one of the premium channels, the cable company's technician inserts the proper IC in the proper converter-unit circuit. Premium channels involve a monthly charge in addition to the basic cable connection fee.

Rooftop Reception of Satellite TV

Many electronic firms offer a complete line of dishes, converters, low-noise amplifiers, and signal processors for home reception of satellite telecasts. The

FCC has given preliminary authorization to several firms who propose to beam their satellite telecasts directly to rooftop antennas and completely bypass the cable companies.

Since the satellite's transponder is low power and its output greatly weakened by the path loss of some 23,000 miles separating the satellite from the earth's surface (about 180 decibels), a very high gain parabolic antenna is needed. In general, the larger its diameter, the greater its gain (and its cost). Heathkit offers a dish 3 meters in diameter (a little less than 10 feet) and fitted with a rugged mounting base designed to hold the antenna solidly in place (Figure 10.10).

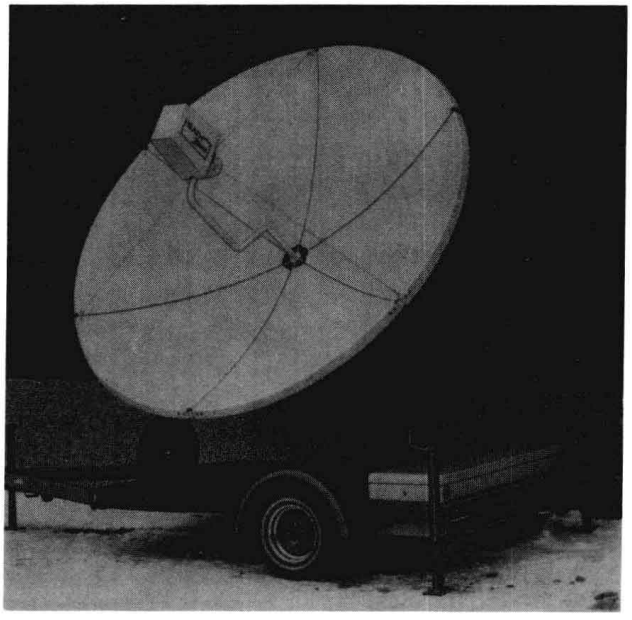


Figure 10.10 Portable High-Gain Antenna Heathkit's 3 meter dish mounted on a trailer permits prospective buyers to try out the equipment at home and determine whether the location is satisfactory for satellite reception. The converter and low-noise amplifier are mounted at the focal point of the parabolic antenna. Output is at 500 megahertz. (Photo courtesy of the Heath Company)

The 4 gigahertz satellite signal (UHF) must be converted to 70 megahertz (VHF) and then amplified at this frequency by a low-noise amplifier (LNA) before being fed to a coaxial transmission line that carries the signal into the home. This unit is mounted at the focal point of the parabolic antenna to keep system losses as low as possible. (Transmission line losses are much less at 70 megahertz.) The coaxial cable inside the house is connected to a satellite receiver (Figure 10.11) that tunes in the various satellites, processes the satellite signal to make it compatible with the TV set's circuits, and finally converts it to some unused channel. Its output connects to the VHF antenna terminals of the TV receiver.

An installation for satellite reception represents a substantial investment. At present all the basic components described sell as a package for around $3500. With such an installation in a good receiving location, it is theoretically possible to pick up over 100 satellite TV signals, but 20 to 30 would be more realistic.

Figure 10.11 Satellite Receiver The Heathkit SRA-8100 satellite receiver connects the output of the 500 megahertz LNA amplifier and the input terminals (usually UHF antenna) of the TV set. Assisted by digital memory, the receiver is capable of 24-channel electronic tuning. Circuitry permits it to turn itself on or off. (Photo courtesy of the Heath Company)

Most of the programs picked up from satellite signals have been produced at great cost with the expectation that revenues derived from their use by cable TV companies would provide a profit for the originating firms. Enjoying these programs in your own home without paying a fee is what the federal courts call *theft of program*. Some programs (e.g., Home Box Office and Showtime) are scrambled before they leave the satellite's transponder unit to prevent unauthorized use. Buying an electronic descrambler to recover the program is technically considered theft of program since no payment is made to the originator of the program.

FUTURE OF RADIO AND TV BROADCASTING

New developments in TV and AM broadcasting will in the immediate future improve the quality of audio transmitted and received. Closest to realization is stereo audio for television. The FCC has already approved several possible stereo technologies, and TV broadcasters uniformly support rapid conversion to stereo sound transmission.

The TV station's sound channel is already FM with a uniform response from 50 to 15,000 hertz. The bottleneck has always been in the receiver, which included a poor audio amplifier (usually a single power transistor) and a completely inadequate loudspeaker 3 or 4 inches in diameter. Under the most favorable conditions the power output was only a few watts, the frequency response perhaps 300 to 3000 hertz, and the distortion level in the range 5% to 7% at best. What was being received was far removed from what was actually being broadcast by the TV station.

Today, with so many high-quality stereo amplifiers in home use, owners see no excuse for poor TV audio, and set manufacturers have determined that a market exists for TV sets with quality audio reproduction. The top lines of a number of set manufacturers in 1983 include receivers with improved audio amplifiers and large, powerful dynamic speakers. Though FM stereo hasn't arrived yet, set designers will be ready when it comes. To transmit FM stereo over an existing FM channel, it is always necessary to include a subcarrier to provide additional information required for the left and right channels. Present TV exciter units do not include provision for subcarrier generation and modulation. However, manufacturers have already developed and field-tested prototypes (early models) with this feature.

Future of AM Stereo

Today more listeners tune to the FM broadcast band (88 to 108 megahertz) than to the AM broadcast band (1600 to 540 kilohertz). Stereo FM broadcasting accounts for listeners' preferring FM, particularly if the broadcasts are channeled through a modern AM/FM tuner, a well-designed stereo amplifier, and two or more wide-range dynamic speakers. By comparison, AM broadcasts sound limited in frequency and dynamic range.

As in the case of TV reception, the problem with AM reception has been largely receivers that included poor audio amplifiers and small, inadequate loudspeakers. AM station managers, wanting to overcome FM's lead in listenership, concluded that AM stereo would have the greatest appeal. The FCC has approved two or more prototype systems and has left it to broadcasters to pick one system as the industry standard. Industry leaders still have not reached agreement on the issue, so the FCC has not yet launched the swing to AM stereo. By 1983 a number of AM stations were transmitting in stereo under special FCC authorization, but AM stereo has not yet captured a large listening audience.

Future of Commercial TV Stations

With some 4400 cable TV systems in operation and a dozen business consortiums clamoring for transponder space on soon-to-be-launched communica-

tions birds, owners of commercial TV stations are beginning to worry about the future. In the early 1980s the audience for commercial TV showed a steady, though not alarming, decrease for two reasons. First was the steady increase both in the number of TV cable systems and in the number of newly manufactured TV receivers that were cable ready—that is, equipped with built-in electronics to receive more than twenty-five channels when the cable was connected to their antenna terminals. Second was the destabilizing effect of the number of TV viewers who purchased the antenna, the frequency converter, the low-noise amplifier, and the tuner-processor necessary for a complete satellite earth station with a potential of more than twenty-four new channels. However, 1984 studies show that audience loss for commercial TV has slowed and, in some areas of the country, has stopped because viewers find local programming superior to cable TV programming.

Current political pressures to deregulate the FCC and to weaken its rules and regulations have also clouded the future of commercial TV. Under present FCC regulations, the commission has complete authority over both TV stations and TV cable systems. The FCC requires cable systems to carry the programs of all major network affiliates in a certain community without competition from out-of-town affiliates. Under deregulation this requirement might be set aside and leave local TV broadcasters with an audience composed largely of viewers in rural or suburban areas where the cable systems find it uneconomical to operate.

AMATEUR RADIO

Since before the first radio legislation in 1912, radio amateurs, or *hams*, have played an important role in the development of telecommunications in the United States. Radio amateurs constructed and operated their own receivers and transmitters, at first without benefit of a license, frequently using initials as identifying call letters. When the Department of Commerce assumed control of radio communications, it required all radio amateurs to be licensed. It also established code tests and written examinations dealing with radio theory and with regulations governing radio communications. A radio amateur was defined as "one who engaged in radio communications because of interest in the communication art and not for monetary reasons."

Applicants who passed the sending and receiving test in international Morse code and the written examination in radio theory and regulations were given an official operator and station license. Call signs were made up of a numeral plus two letters. A typical call sign would be 7SY. The number identified the amateur station as being located in the Pacific Northwest; the SY identified the particular station. The same procedure for licensing radio amateurs prevails today.

Contributions of the Amateur Radio Service

Amateur radio operators have made important public service contributions in providing emergency communications during disasters. When other communications systems have been disrupted by floods, earthquakes, or hurricanes, amateur radio has frequently provided the only link with the outside world. Operators have carried requests for medical assistance, ordered needed supplies, and reassured anxious relatives. In wartime the armed forces rely heavily on the pool of radio operators for emergency communications services.

Amateur radio has also contributed to the growth of electronics. Almost all the breakthroughs in the field were made by licensed radio amateurs. A large proportion of the chief engineers of the first radio and TV stations acquired their first electronic training as ham operators. Many business executives whose only experience with electronics was acquired in operating amateur stations now head some of the nation's most prestigious electronics firms.

Amateur Radio Today

For two reasons the trend is for radio amateurs to do more operating of their station equipment and

This amateur radio station shows comparatively little equipment in use but provides SSB and AM phone and CW (code) on all amateur bands between 1.8 and 30.0 megahertz with 200 watts. Accessory meters atop the solid-state transceiver read RF power output (left) and SWR, or standing wave ratio (right).

less experimentation. First, with the swing to integrated circuits (ICs), individual constructional components are difficult to obtain without long delays, and ICs are beyond the capabilities of many hams. More important, amateur radio is increasingly becoming an avocational interest—an escape from stress and tension. Most equipment in service today is factory built but used by individuals with a technical understanding of its operation.

Unlike citizens band operators, who are limited to 4 watts of power on forty crowded channels centered around 27.0 megahertz, amateur operators may use up to 1000 watts of power on a large number of frequencies extending from near the AM broadcast band (1.8 megahertz) in harmonically related bands, each one having twice the frequency of its low-frequency counterpart. The old terminology is still in wide use: 4.0 megahertz is referred to as 80 meters, for example.

Today amateur operation is not restricted to AM phone and international Morse code (known as CW). FM is in wide use in the higher frequency bands. Slow-scan television allows operators to see each other while talking. Probably the newest and most popular mode of communication is called **SSB,** or **single-sideband transmission,** which is actually a form of AM but has more signal punch. On the lower-frequency bands, particularly, you can tune in to roundtables (four or five stations talking with each other in rotation) or nets (groups of twenty or thirty stations practicing the handling of message traffic, ready for an emergency to develop). On the higher frequencies are **DX operators,** who strive to work other amateur stations in foreign countries. Foreign stations contacted by DX operators are licensed much as our own and operate in the same frequency assignments by international agreement. There are over 1 million licensed amateur radio operators in the world, 342,000 in the United States.

OSCAR Space Satellites

Since 1961 hams have had one or more satellites aloft and in communication with amateur stations on the ground. These satellites were built by amateurs. Costs were underwritten by other amateurs. The satellites are known as OSCARs (an acronym for orbiting satellite-carrying amateur radio).

OSCARs 6 and 7 are currently orbiting the earth in 900-mile-high polar orbits, serving as a space radio relay for the amateurs of many countries. Each satellite was launched by NASA as part of a payload destined for outer space. OSCAR satellites incorporate a transponder unit that receives code, voice, and even slow-scan TV signals on one frequency and retransmits them on a different frequency. A program for their use in public schools is under way. They will be used in teaching space science and physics.

American Radio Relay League

Nearly half of the amateur operators in the United States belong to a nonprofit organization that acts as their advocate before the FCC and in international meetings where frequency assignments are decided. Known as the American Radio Relay League, it publishes *QST*, (a monthly magazine) and many paperbacks. Publications include an annually updated *Radio Amateur's Handbook*, a license manual to aid prospective hams in passing FCC exams, and many

special-interest booklets. The ARRL's address is 225 Main Street, Newington, Connecticut 06111.

The league's club and training department can answer questions about volunteer license examiners, licensing, and putting on amateur radio exhibits. The league also offers some excellent promotional and training films. Canadian amateurs are represented by the Canadian Radio Relay League (CRRL), a self-governing division of the ARRL.

CITIZENS RADIO SERVICE

The FCC has recently reclassified both the Amateur Radio Service and the Citizens Radio Service into a new category of telecommunications known as Personal Radio Services. Beyond being included in one common category, there is little similarity between the two groups.

The Citizens Radio Service was originally conceived by the FCC as a communications service completely different from the hobbying encountered today in the spectrum around 27.0 megahertz. The citizens band (CB) was intended primarily for essential business or personal short-range radio communication. It was also to provide a facility for signaling and for radio control of objects and devices.

The FCC also envisioned the use of Citizens Radio Service to provide communication to isolated farms, ranches, and recreation areas. Strictly forbidden were idle conversations with other CB operators. Also off-limits was the working of distant stations that during periods of strong sunspot activity, "skipped in" by multiple reflection off the ionosphere. Power output of the AM transmitter was limited to 4 watts on any one of twenty-three crystal-controlled channels. Only FCC-approved equipment could be used. No license examination was required. Participants needed only an application.

Problems Facing CB Radio

The Citizens Radio Service has moved in an entirely different direction than originally conceived. For one thing, the FCC greatly underestimated the number of citizens who would apply for a CB license. Licensing CB operators quickly became the largest single service the FCC administered. As of December 1977, the FCC has issued more than 12 million station licenses. The number of sets in service was estimated at around 25 million. That is, about 17.6% of the nation's homes had CB installations, and millions of unlicensed CB sets were in operation.

Enforcing FCC regulations with just a few personnel in the field has proven impossible. Regulations against idle conversations and the working of distant stations during periods of skip have been widely ignored. The power limitation of 4 watts output has been violated by enthusiasts who have installed 1000 watt amplifiers, and malicious interference has been widespread. Courtesy between operators on the same channels is conspicuously absent. The CB allocation has become tremendously overcrowded with hobbyists more interested in gossiping than in communicating.

However, though the number of stations has continued to grow (the total is now estimated at 30 to 40 million), fewer signals seem to be on the air. Also, the signals on certain channels are cleaner and have more punch because they use single-sideband transmission (SSB), a specialized form of amplitude modulation. As has been the case for many years, many of the AM signals sound fuzzy and distorted—the result of overloading electronic circuits by using a power microphone to squeeze out just another watt or two of signal.

The FCC is struggling to maintain some control over CB operations. It hopes to move all operators to the UHF spectrum around 900 megahertz, but the services presently operating on these frequencies must first be moved to new channels. In April 1983 the FCC announced that it was abandoning the CB licensing procedure because it was largely disregarded by CB operators and not worth the $400,000 a year in operating costs. However, the FCC will continue to monitor CB frequencies to ensure that all operators comply with the rules governing radio communication.

On the positive side, CB has made a real public service contribution when national or regional communications emergencies have occurred. Since many

A citizens band operator reports a traffic accident with his mobile installation. New FCC rules permit CB operators to use their equipment for public service. Channel 9 is usually reserved for emergency traffic reports. (Photo courtesy of E.F. Johnson Company)

CB operators have equipment installed in their cars, they can quickly report a traffic accident or summon emergency help, particularly now that channel 9 of the forty channels currently assigned is considered throughout the country as the emergency channel and is used primarily for reporting emergencies.

Radio Emergency Associated Citizens Teams

Although CBers have no national organization with political clout to represent them before the FCC, they do have radio emergency associated citizens teams (REACT). Founded in 1950 by Hallicrafters, a manufacturer of electronic communications equipment, the organization has been sponsored in recent years by the research and development department of General Motors primarily to provide emergency assistance to motorists. REACT members continuously monitor channel 9 so that they can immediately respond to any radio request from motorists in trouble. In many cases REACT members have been known to provide emergency gasoline, to assist in changing a tire, and to relay requests for medical assistance.

CB as a Training Ground for Amateur Operators

The 4 watt power limitation, the forty channels, and the intolerable level of interference convince many CB operators that the amateur bands are more satisfactory. Unlike the CB channels, the amateur bands have many frequency allocations, a maximum power of 1000 watts, and the potential for contacting stations all over the world instead of just a few blocks away. As a result, many CB operators learn the code, study theory and FCC rules and regulations, and achieve amateur status. Surprisingly, a number of amateurs have also obtained CB licenses and have installed mobile units in their cars, principally to give themselves access to emergency channel 9 and its potential for roadside service.

SUMMARY

Telecommunications has come to mean what radio pioneers referred to as long distance communication. Technically, it includes communication over metallic wires, communication through use of electrical impulses (telegraph and telephone), and communication by electromagnetic fields, either through free space (radio and TV broadcasting) or over coaxial cable (cable television).

Radio communication began shortly after 1900, but widespread interest in radio broadcasting did not develop until shortly after 1920. As the number of radio stations rapidly increased, Congress appointed regulatory agencies to deal with equitable assignment of frequencies, technical standards of performance, and the licensing of operating personnel. In 1934 Congress created the Federal Communications Commission, which is in control of telecommunications today, though its authority has been somewhat diminished by deregulation.

Today, full-color TV has become a reality, and the FCC has approved the concept of stereo audio

for TV stations. Commercial TV stations have a new competitor: a number of cable TV systems that offer perhaps thirty different programs originating in a distant production center and brought to the system's distribution equipment by a satellite UHF relay station. Cable TV systems will soon be joined by a new service that proposes to beam TV programs directly from the satellite to the rooftop antenna of a paying subscriber.

FM, because of its appeal to stereophonic music enthusiasts, has become the dominant force in radio broadcasting. AM station licensees are trying to compete, but AM stereo standards are still in a state of flux and reliable AM stereo equipment in short supply.

Amateur radio has won a number of political decisions that will benefit many operators. Much of the pioneer 160 meter band has been freed of strong interference caused by LORAN stations using the same spectrum. A new 30 meter band has been made available for code operation. The FCC also has denied a petition for abolishing the code test in all amateur operator examinations.

The Citizens Radio Service (CB) has become less popular since 1980, with less band activity and slow sales of equipment. The principal purchasers of new transceivers are motorists who like the security of being able to reach emergency channel 9.

KEY WORDS AND CONCEPTS

Amateur Radio Service
amplitude modulation (AM)
Citizens Radio Service (CB)
DX operator
field
frame
frequency modulation (FM)
iconoscope
interlaced scan
intermediate frequency (IF)
monochrome
mosiac
single-sideband transmission (SSB)
telecommunications
transponder
wireless

SELF-TEST QUESTIONS

1. How does the authority of the Federal Communications Commission differ from that of the Federal Radio Commission, the regulatory agency it replaced?
2. Why are communications systems using amplitude modulation (AM) particularly subject to electrical interference?
3. In AM transmissions, the amplitude, or power, of the radio frequency carrier wave is varied at a voice or music frequency rate. In FM transmission, what other characteristic of the carrier wave is varied?
4. What new development in FM transmission and reception led to a rapid expansion in the number of FM stations and FM receivers?
5. How did Edison's early experiments with motion pictures help television pioneers define the problems of TV picture transmission and suggest some possible solutions?
6. Why was an all-electronic television scanning

system much to be preferred over Nipkow's system, which used whirling mechanical discs?

7. Zworykin's iconoscope, the first TV camera capable of high-quality picture reproduction, was a highly specialized transducer. What form of energy did it convert to pulses of DC current?
8. Explain briefly how TV transmission and reception in the United States uses both AM and FM technology.
9. Color television is based on the transmission of what three primary colors?
10. Originally, cable television systems served only to relay distant TV signals to local viewers. What are some of the additional services currently offered by cable TV systems?
11. Why are all licensed radio amateurs competent in electronic theory and in the sending and receiving of messages in international Morse code?
12. How does the maximum power permitted amateur stations compare with the 4 watt limitation imposed on CB stations?
13. How does the present Citizens Radio Service differ from the FCC's original concept?
14. What special use is made of channel 9 by CB operators across the country?

11 Recording and Audio Systems

Rosario Bourdon and members of the Victor Salon Orchestra make a microphone recording in the mid-1920s. (Photo courtesy of RCA)

Recording and Reproduction of Sound
- Edison Phonograph
- Edison's Recording Unit
- Record Discs
- Sound-Motion Pictures
- Electronics and the Recording Process
- Union of Radio and Phonograph
- Improvements in Record Discs
- Record Disc Speeds

Magnetic Recording and Playback
- Earliest Tape Recorders
- How Magnetic Tape Recording Works
- Improvements in the Basic Unit

Audio System Components
- Amplifiers
- Loudspeakers
- Pickup Cartridges and Styli
- Microphones
- AM/FM Tuners
- Other Accessory Equipment
- Room Acoustics

Developments in Recording and Playback
- New Coatings for Tapes

Cassettes
Digital Technology

Summary

Key Words and Concepts

Self-Test Questions

Laboratory efforts to record the human voice have been under way at least since the Civil War, and the first successful acoustic recording was made and demonstrated a century ago. Interest in the improvement of its technology was slowed until the first decade of the present century by the lack of a device that would amplify a weak electrical signal into a more powerful one. Lee De Forest invented such a device: the audion, or three-element vacuum tube. However, development of electronic circuitry that would use this new active device effectively in the recording and reproduction of voice and music required another twenty years. The invention of the transistor in 1947, although hailed as a breakthrough, also required a long transitional period before it could be used effectively.

Much of the technology developed in the recording of aural information proved valuable when the need arose to develop techniques and equipment for the recording and playback of visual information, not only in the field of magnetic tape but also in that of video disc recording. Even today, with visual recording and playback equipment under continuous development, this specialized area of electronics owes many of its basic concepts and skills to those developed in our much longer and more extensive experience with aural recording.

In this chapter we first discuss the development of the technology and equipment for the recording of voice and music. Today this field is simply referred to as audio after the Latin word *audire*, which means "to hear." We then turn to a discussion of audio systems, usually called home music systems. We describe the individual components of such a system and the desirable characteristics of each.

RECORDING AND REPRODUCTION OF SOUND

The first equipment capable of recording the human voice and playing it back was a result of Thomas Edison's seeking a device that would record the characters of Morse code for later decoding into letters and numbers. In 1877 he decided upon a new recording device consisting of a metal stylus (pointed cutting tool) attached to the telegraph sounder. In moving up and down with each electrical impulse, the stylus would emboss the code's dots and dashes on a wax-coated paper tape. The tape, fed from a reel, was advanced by a clockwork motor. The unit worked perfectly with little change in the original design and was manufactured and sold by Edison's company for years.

Impressed by the variable embossing effect of the stylus on the wax-surfaced tape, Edison took a diaphragm from an experimental telephone and attached it to a similar embossing stylus. Holding the unit in contact with the moving tape, he shouted into the cupped diaphragm. Winding the tape back, he held the stylus in the embossed groove and listened. Something could be heard but not understood. Edison then commissioned a mechanic to make a revolving cylinder of metal turned by a hand crank and equipped with a threaded screw to move th recording stylus and its attached telephone diaphragm laterally from one end of the drum to the other.

Placing a piece of tinfoil carefully over the metal cylinder and securing it with elastic bands, Edison positioned the stylus and started turning the crank. Into the mouthpiece he recited the nursery rhyme "Mary Had a Little Lamb." He then repositioned the stylus and again started cranking. Although the diaphragm's output was very tinny and distorted, the words could be distinguished. For the first time the human voice had been successfully recorded. Edison

quickly patented his device, for Alexander Graham Bell was pursuing the same goal with the same basic type of equipment. (Bell's recorder was the forerunner of the device known today as the Dictaphone.)

Edison Phonograph

Edison turned his basic recording-playback device over to his staff for development and production. The result was a remarkably sophisticated instrument for its time (Figure 11.1). The recording was made on a wax cylinder about 4 inches in length and 2 inches in diameter. The cylinder, driven by a spring-wound motor, turned at a constant speed of 100 rpm (revolutions per minute). A diamond-tipped stylus attached to a flexible diaphragm (usually made of mica) was joined to a flaring metal horn. Its movement across the record's surface was assisted by a lathe screw and guide that moved the stylus rather than the grooves alone. The mechanical work throughout was of the highest quality. Even today, some of the better Edison cylinder recordings sound remarkably clear.

Edison's Recording Unit

Recordings were made on a unit heavier and larger than the Edison phonograph unit but otherwise much the same. The horn had to be much larger physically so that the performers could stand directly in front of it to produce maximum volume on the recording. The sound waves, traveling down the horn, vibrated the diaphragm and attached stylus, which rose and fell in a pregrooved wax cylinder. This type of vertical recording was known as **hill-and-dale recording.** Edison had also patented, but did not use, **lateral recording,** in which the stylus moves from side to side in a groove of constant depth.

Record Discs

Edison's sharpest competition came from German-educated Emile Berliner, who had come to America to work for the Bell Telephone Company as a researcher. Berliner's gramophone, as he called it, broke entirely with the Edison tradition of cylindrical records (Figure 11.2). Berliner introduced the so-called pancake record, which was about 10 inches in diameter and three or four times as thick as modern vinyl discs. It turned at about 80 rpm and used lateral

Figure 11.1 Edison Gem Phonograph Manufactured in 1912 by the National Phonograph Company of New York, this phonograph was fairly sophisticated for its time. A lathe screw advanced the diamond stylus across the record cylinder. (Photo courtesy of National Museum of American History, Smithsonian Institution)

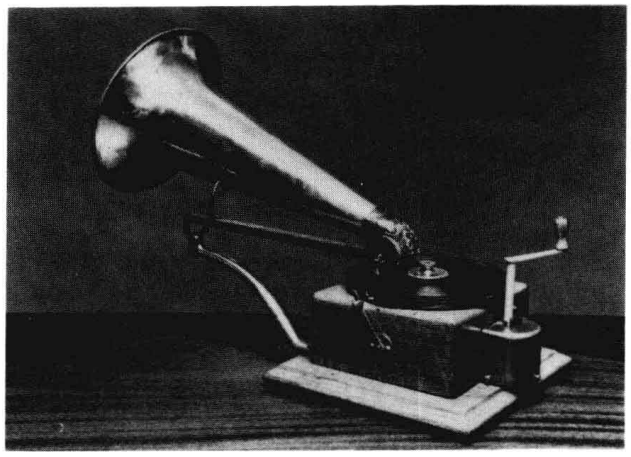


Figure 11.2 Model of the Original Berliner Gramophone This improved model used a spring-wound motor and became the basic design of all later phonographs. The machine played disc records laterally recorded. (Photo courtesy of National Museum of American History, Smithsonian Institution)

(side-to-side) modulation of the groove with unvarying depth instead of hill-and-dale recording. Berliner also introduced a mastering process whereby a metal stamper was made by electroplating from the original recording. The stamper was then used to stamp out countless duplicates from lumps of plasticized shellac softened by steam. Record companies in this country employ the same production process today.

Sound-Motion Pictures

Late in 1982, motion picture film archivists verified that Thomas Edison had apparently been the first person to develop sound-motion pictures by teaming the motion picture system he had invented with speech and sound effects from his recorded disc. (The discs and films were accidentally separated and remained stored in separate media archives until 1982.) Picture and sound synchronization was maintained by mechanical linkage between the motion picture projector and the phonograph turntable. Although the year of production of these early sound-on-disc films was 1912 or 1913, the quality of reproduction and the synchronization between film action and sound were good.

Electronics and the Recording Process

In 1925, without a word of advance publicity, the Victor Talking Machine Company started shipping its distributors a new series of 10 inch and 12 inch records, all of which were recorded by the Western Electric Process. That is, the voices and instruments of the performers were picked up by a microphone and amplified by a vacuum-tube amplifier. The output of the amplifier was then used to energize a magnetic recording head. Since the magnetic recording head had almost unlimited cutting power compared to the acoustic head it replaced, it was possible to cut the record grooves much more heavily than before.

The immediate effect was a great increase in the volume obtained when the record was played back. A more subtle difference was the improved high-frequency response of the discs. The sibilants in human speech (the *s* and *sh* sounds) were easily distinguished, and the high-frequency tones of the violin sounded relatively natural for the first time. But a major roadblock still stood in the way of better audio quality. The recordings were vastly improved by the use of electronics, but the playback equipment was still strictly acoustic. Moreover, it distorted badly on loud musical passages, and with no control over the volume, many listeners resorted to stuffing a rag down the neck of the horn to make the sound less strident.

Developments in radio receiver circuitry to match improvements in the quality of radio broadcasting were proceeding at a rapid pace. The introduction of AC power supplies in 1926–1927 made possible the use of larger vacuum tubes with much larger power output. The vacuum tubes, in turn, made practical the use of the dynamic speaker, which repro-

duced both voice and music at high levels with a minimum of signal distortion. Radio, the newcomer, was thus achieving far better audio quality reproduction than the phonograph industry had been able to develop in nearly fifty years.

Union of Radio and Phonograph

The next logical step was the addition of a phonograph turntable to a radio receiver. This combination immediately became known as a radio-phonograph. A few technical matters had to be resolved before the device could be widely sold. First, a transducer had to be designed to transform the side-to-side movement of the stylus in the record groove to corresponding changes in an electric current. This device was called a **phonograph pickup.** Converting its electrical output back to sound waves required an audio frequency amplifier to build up the strength of the electrical signal, and then another transducer to change the electrical waves back to sound waves. This transducer was the radio's loudspeaker.

Since the audio frequency amplifier, power supply, and associated loudspeaker were already incorporated in the radio receiver, it was only necessary to include a simple switch that connected its input circuit alternately to the radio receiver's detector tube or to the leads from the phonograph pickup. At the same time the conventional spring-wound turn table was replaced with one driven by an electric motor. The audio quality of records heard over the new radio-phonograph was generally far superior to that of even the best acoustic phonographs. Today's AM/FM receivers with their high-quality audio systems use the same basic interconnection of radio and phono units.

Improvements in Record Discs

The inherent abrasiveness of the shellac 78 rpm records quickly changed the carefully shaped playback stylus into a sharp-edged cutting tool that soon destroyed the recording groove. For example, a typical RCA Victor radio-phonograph would use Victor Red Seal steel needles, usually purchased 100 at a time. After every two selections played, it was necessary to stop the machine, remove and discard the needle, and insert a new one. This routine continued until the introduction of the long-playing (LP) record in about 1950. Made of pure vinyl resin, it was almost free of any abrasive qualities and made possible the use of semipermanent sapphire and diamond styli.

Record Disc Speeds

In the 1950s Dr. Peter Goldmark, then president and director of research of CBS Laboratories, began to work on a new recording system that would provide playing time for an entire symphony or concerto (40 to 50 minutes) on a single 12 inch vinyl disc. As an engineer, he recognized that such a system would require a much smaller recording groove (1 mil instead of 3 mils), a much slower turning speed (33 1/3 rpm instead of 78 rpm), and a boosting of the high audio frequencies during the recording process and of the low audio frequencies during playback. The boosts were necessary to preserve the high-frequency response of the disc, which would otherwise be lost because of the low rotational speed.

The LP record won immediate acceptance, but finding equipment capable of playing it proved troublesome since no 33 1/3 rpm speed was available on home record players, and the existing pickup heads were far too heavy to use with a 1 mil stylus. To add to the confusion, RCA Victor refused to go along with the new Columbia Records development and announced a new record of its own. It was 7 inches in diameter, turned at 45 rpm, used a 1 mil stylus, and could be played only on a special RCA Victor record player. It was the forerunner of the millions of pop recordings sold today.

The dilemma of the average radio broadcast station of this period is shown in the photo of the radio station control room taken about 1952 (Figure 11.3). Note the two turntables, each originally with two speeds—one of 78 rpm for standard shellac records, the other of 33 1/3 rpm for the 16 inch electrical transcriptions in wide use by radio stations at the

Figure 11.3 Radio Station Control Room This photo was taken at station KRVM-FM about 1952. Note the 16 inch electrical transcription on the left foreground turntable and the LP on the turntable to right. Separate pickup arms were required for regular 78 rpm shellac discs and for the new LP vinyl pressings.

time. Note that the original heavy pickup arms are still in service for shellac records and transcriptions, while new, lighter arms are provided for the LP records. The 45 rpm discs could not be played at all.

MAGNETIC RECORDING AND PLAYBACK

Experimentation with magnetic recording of sound has been going on since 1898, when Valdemar Poulsen, a Danish scientist, discovered the basic principles involved. By passing a steel wire in front of an electromagnet through which the pulsating current of a telephone was being led, he found that a permanent magnetic pattern of the human voice could be impressed on the wire. To play it back, he passed the same wire over the same electromagnet while listening to the earphone.

Although Poulsen's wire recorder worked quite well, it was plagued with mechanical difficulties. The tempered steel wire was difficult to handle, it tangled easily, and was difficult to repair when it broke. Worst of all, it quickly cut a groove into the record-playback head, necessitating its constant replacement. Poulsen lacked a vacuum-tube amplifier to increase the output level, but it hadn't yet been invented. After World War II both RCA and Webster-Chicago manufactured wire recorders, but because of constant tangling and breaking of the wire, they were never popular (Figure 11.4).

Earliest Tape Recorders

In the closing days of World War II, American troops captured a small group of German noncombatants equipped with two audio recorders of a new type.

Figure 11.4 Webster-Chicago Wire Recorder of the 1940s Although some broadcasters used wire recorders after World War II, the wire tangled readily and was difficult to splice.

Instead of using a recording lathe, a powerful motor, and a high-power amplifier to drive a cutting stylus, they used reels of paper tape coated with iron oxide. A well-designed clockwork motor powered the take-up and supply reels. A simple, battery-powered audio amplifier provided excellent-quality reproduction.

Recognizing the importance of this find, technical personnel sent the units to the Signal Corps labs in Washington, D.C., for study. The recorders had been manufactured in a small factory near Munich, and additional data were obtained from the designers at the close of hostilities. These two machines became the prototypes of every tape recorder, tape deck, and tape player built since that time. They even led the way, eventually, to the technique of video recording.

How Magnetic Tape Recording Works

The output of a high-quality audio amplifier in the form of an AC current, varying at a voice or music rate, is led through the windings of a specially designed electromagnet that has a very small gap between its north and south poles. The varying current in the windings gives rise to a varying magnetic field, which is most intense in the area of the gap in the electromagnet. The electromagnet has been lapped and polished so that the coated tape can glide over it readily.

As the tape moves at a constant speed through the varying magnetic field, the molecules of its coating (typically iron oxide) are rearranged into a magnetic pattern corresponding to the voice or music currents flowing in the windings of the **recording head.** This process of rearrangement into a new magnetic pattern takes place more readily if the tape coating is at the same time subject to a low-frequency radio frequency field produced by a transistor in an oscillatory circuit. It is usually fix-tuned to about 50 kilohertz.

When the tape is rewound on the supply reel, it is ready for playback once a switch is turned. The switch disconnects the recording head from the amplifier's output and connects it to one of the unit's inputs. At the same time a loudspeaker is connected to the amplifier's output circuit. As the tape moves across the face of the electromagnet, its varying magnetic pattern sets up a weak alternating current in its windings that varies at a voice or music rate. This AC is then strengthened by the amplifier into a powerful signal capable of operating the loudspeaker at high volume. The resulting sound waves are essentially an exact reproduction of the voice or music originally picked up by the tape recorder's microphone.

Since the first machines were produced in the United States in the early 1950s, the industry standard for the tape speed of reel-to-reel machines has been 7 1/2 inches per second. Therefore, it is possible to play a tape recorded on one brand of equipment on a machine of a different manufacturer. Unfortunately, standardization did not take place in videotape recording.

Improvements in the Basic Unit

While the basic recorder-playback unit can provide excellent performance, it can be enhanced by using separate heads for recording and playback functions. Each head is designed for maximum effectiveness. To erase an existing tape recording, a separate **erase head** is also provided. It is energized by the same 50 kilohertz oscillator used to assist in the recording process, and its switching is so arranged that the chance of accidental erasure of a recording is minimized. Today's tape units are ideally suited to the recording and playback of stereo programming.

No modern tape deck would be considered complete without the addition of a meter to monitor the recording level during the recording process. Using the meter, the operator can maintain recording level at the optimum value. Known as a **volume unit,** or **VU meter,** it is essentially an AC voltmeter with special characteristics. It is connected across the amplifier's output circuit.

AUDIO SYSTEM COMPONENTS

In this section, we examine individual electronic components such as amplifiers, microphones, phonograph pickups, and AM/FM tuners. We also examine how each component contributes to the operation of a total sound system typically found in the home.

Amplifiers

The modern audio amplifier is not markedly different from the vacuum-tube amplifier of the same power rating produced twenty-five years ago. It consists of a preamplifier, which incorporates a number of inputs for various program sources and a number of controls, and a power amplifier, which drives the loudspeakers. Because two comparatively small cylinders in a balanced circuit provide a much smoother flow of engine power than a single, very large cylinder, the amplifier uses two heavy-duty transistors in the output stage in a balanced circuit called a **push-pull circuit** (Figure 11.5). One transistor conducts heavily over half of the signal cycle while its partner idles. Then the second transistor conducts heavily over the next half of the cycle while the first transistor idles. A push-pull connection of the transistors in the output stage produces more than twice the output of a single transistor with but a fraction of the signal distortion that one large transistor of the same power would create.

The amplifier we have described is of the **monophonic,** or mono, type and is still manufactured and sold for public address and paging systems and for electronic musical instruments such as organs. In these applications a single sound channel is usually adequate.

The last decade has seen a massive move from monophonic to **stereophonic** recording and reproduction of sound. Virtually all albums offered for sale are stereophonic, as are nearly all tape cassettes and reel-to-reel tapes. (The trend is away from multitrack cartridge tapes to the tape cassette with its simpler mechanics and higher-quality reproduction.)

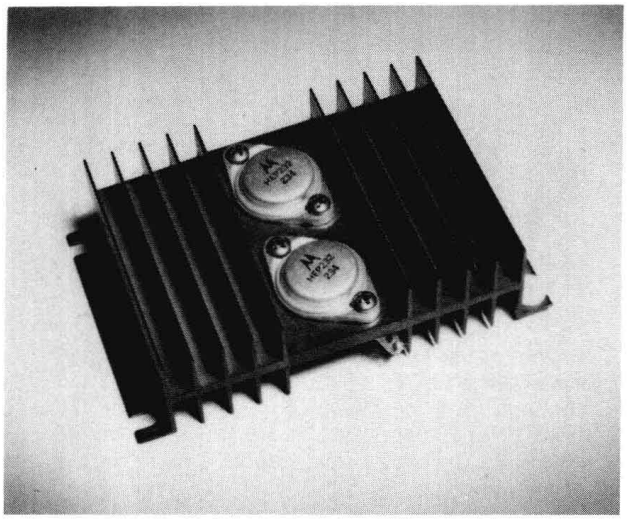
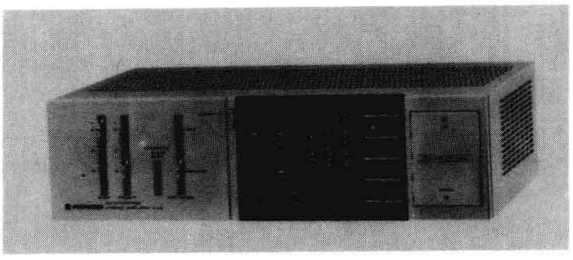


Figure 11.5 Power-Output Transistors Connected in a Push-Pull Circuit For increased output and reduced distortion, these Motorola transistors are connected in a push-pull circuit and mounted on a finned heat sink. Heat generated by the transistors flows away from their surfaces and is dissipated into the surrounding air to permit them to operate at a safe temperature.

Stereo Reproduction

The purpose of stereo reproduction is to create a spatial effect—that is, to add another dimension to recorded music. Stereo reproduction conveys a more vivid sense of reality and creates the impression that the listener is actually seated in the studio. The illusion of dimension is achieved largely by using a minimum of two microphones, each feeding a completely independent recording amplifier. In turn, the amplifiers feed completely independent recording heads, which lay down two separate sound tracks on a single recording tape. From this master tape



The Pioneer Electronics model A-X8 integrated stereo amplifier couples small size with large-scale audio performance of 65 watts per channel. Frequency response is essentially flat from 20 to 20,000 hertz, and distortion is well below 1%. Inputs are provided for microphone, phono, and tape player. (Photo courtesy of Pioneer Electronics, USA, Inc.)

both stereo tapes can be copied (dubbed) or stereo records cut (mastered).

Using these aural materials effectively to recreate a sense of spatial realism requires that the reproducing amplifier be two completely independent amplifying systems that feed two completely independent speaker systems. However, they can share a common power supply and chassis. With the extreme miniaturization that the use of transistors and integrated circuits makes possible, the total space required is very small. Both the input and output circuits of each amplifier are plainly marked "left" and "right" to correspond with the two microphones used in the recording process—one positioned to the left and one to the right of the performers. Today it is commcn practice to use a number of microphones rather than just two.

Fidelity and Distortion

Since the 1930s, trade magazines have made wide use of the term **high fidelity,** which means faithfulness of reproduction. A high fidelity amplifier, then, is one that exhibits remarkable accuracy of reproduction. What you hear from the loudspeaker sounds essentially the same as the original performance in the recording studio.

Distortion—that is, any change in the audio waveform (except an increase in amplitude) as it passes through an audio amplifier from input to output—interferes with fidelity. The most disturbing form of distortion in an audio amplifier is caused by the improper operation of the equipment's active devices (vacuum tubes, transistors, and ICs). When active devices operate with incorrect values of voltage and current or are overloaded by the signal, they tend to create new and additional frequencies as high as five times the fundamental frequencies contained in the original audio input's waveform. Addition of these frequencies to the original signal may cause it to sound harsh and discordant. A person with musical training can usually detect distortion levels when they exceed 5%. Other people will easily recognize the harshness in sound that 10% harmonic distortion produces.

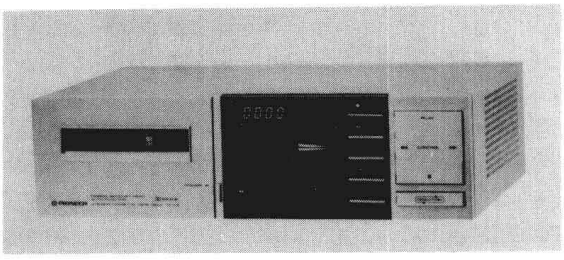
Equalization

In disc recording and playback, the high-frequency response falls off markedly in the inner grooves of the record and improves markedly as the outermost grooves are reached. The improvement is due to the increase in the stylus's writing speed as the outer grooves are approached. Therefore, it is necessary to boost the high frequencies in the recording process and roll off the highs on playback to produce a substantially flat, or uniform, frequency response.

This technology is called **equalization.** In its simplest form it might require only a single capacitor of a specific size and a series resistor to limit the capacitor's action. Equalization greatly reduces both stylus groove noise in disc recording and tape hiss in tape recording, as well as stray pops and clicks of a high-frequency nature.

Noise Reduction Circuits

In the 1980s few advanced tape decks, particularly cassettes, are considered complete unless they



The Pioneer Electronics model CT-X9 cassette player offers autoreverse for uninterrupted music, Dolby C noise reduction, and music search for quick location of recorded selections. A microprocessor makes possible one-touch computerized controls for many operator-selected functions. (Photo courtesy of Pioneer Electronics, USA, Inc.)

incorporate an automatic noise reduction circuit. The principle involved is simple. At normal program levels the program content masks annoying hisses, pops, and clicks. When the program level drops, these noises become evident but can be reduced by high-frequency rolloff. A special sensing circuit of the noise reduction unit monitors the program level, introduces rolloff only during the low-volume passages, and then restores normal, flat response.

Development of Amplifier Standards

During the early 1940s audio amplifiers with a frequency response ranging from perhaps 100 hertz to 10 kilohertz and with harmonic distortion not exceeding 10% at rated output were frequently described in the trade media as "high fidelity." Amplifiers produced by prestigious manufacturers comfortably exceeded these minimal specifications.

As FM broadcasting began in the new 88 to 108 megahertz allocation in the late 1940s, FCC standards for the audio response range of transmitting equipment (50 to 15,000 hertz with not more than 3% distortion at any frequency) became the goal for audio equipment manufacturers. They soon realized their goal because of improved output transformers and increased knowledge of how to use negative feedback to reduce signal distortion and flatten frequency response.

By the mid-1970s engineers had made remarkable progress in producing a new generation of audio amplifiers that employed bipolar transistors to produce tremendous amounts of audio power with amazingly flat response levels and a minimum of distortion. New designs that eliminated the output transformer and used vast amounts of negative feedback were largely responsible for the breakthrough. In the mono and stereo amplifiers of the mid-1980s, there is no absolute standard. Power output is frequently 50 to 100 watts; frequency response commonly exceeds 20 to 20,000 hertz and is flat within a decibel or two; and distortion is well below 1%.

Response of the Human Ear

The human ear, which the various types of stereo amplifiers are designed to affect, behaves logarithmically in its response. In other words, doubling the sound pressure impinging on the ear's diaphragm does not result in a perception of the sound as twice as loud. Sound pressure would have to be increased four or five times for a noise to sound appreciably louder. The ossicles of the inner ear act as an automatic volume control. For this reason, there is little point in producing an extremely powerful amplifier.

Loudspeakers

A **loudspeaker** may be considered a special form of electric motor in which the moving coil reciprocates rather than rotates. As shown in Figure 11.6, it is made up of three basic components: a permanent magnet of great strength with a circular slot to accept the voice coil; a voice coil, or armature, made up of comparatively few turns of fine wire wound over a form; and a paper or fiber cone that may be from 3 to 15 inches in diameter. One end of the voice coil is cemented to the apex of the cone, while the

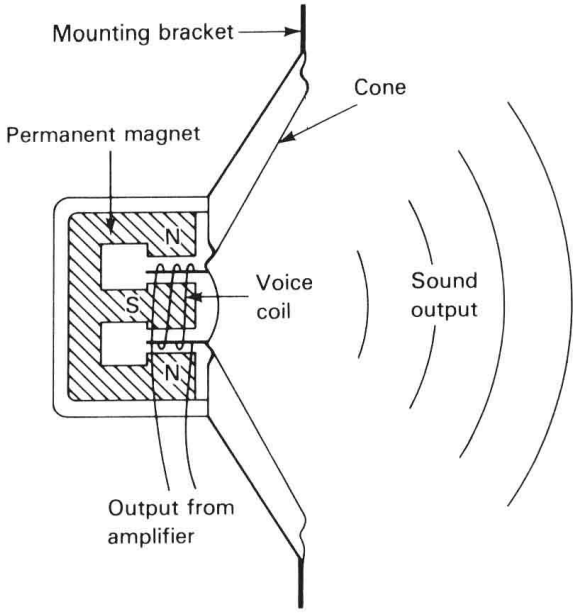


Figure 11.6 Cross Section of a Permanent-Magnet Dynamic Speaker When an AC current, varying at a voice or music frequency rate, flows through the voice coil, its changing magnetic field reacts with the fixed field of the permanent magnet and causes the voice coil and attached cone to move back and forth to create sound waves.

edge of the cone is glued to a ring-shaped flange of fiber attached to the loudspeaker's metal supporting frame. The voice coil leads connect to the secondary winding of the output transformer.

Loudspeaker Construction

When an audio signal passes through the amplifier, the output transformer converts the pulsating DC that the output transistors produce into a low-voltage, high-current AC waveform in the secondary windings. This waveform is led into the voice coil winding. As the current in the voice coil changes polarity, so does that of the magnetic field surrounding it. As a result, the voice coil moves back and forth a short distance as it is alternately attracted and repelled by the magnetic field of the powerful permanent magnet. In moving, the coil also moves the speaker's cone back and forth, to set up waves in the air acted on by the cone. These waves are called **sound waves.** The conversion efficiency from electrical to sound energy is very low—always less than 15% and usually much lower. A typical bookshelf speaker and enclosure may have a conversion efficiency of only 5% to 6%.

Multispeaker Installations

Although each channel of a stereo system requires at least one loudspeaker in a suitable cabinet, it is good practice to use two or three. The largest, usually equipped with a cone about 12 inches in diameter, is specially designed to handle the low frequencies effectively. Another, perhaps with a 6 or 8 inch cone, is designed to radiate the mid-range frequencies efficiently. A very small unit with a cone only 2 or 3 inches in diameter radiates extreme high frequencies very effectively. All three loudspeakers are connected electrically in parallel. A typical loudspeaker system is shown in Figure 11.7.

The cabinet, or enclosure, greatly improves the operation of the mid-range and bass-range speaker but has little effect on the treble-range speaker because the *baffle*, as it is called, prevents interaction between compression and rarefaction waves from the front surface of the cone and from the rear surface of the cone. Interaction could cause near-cancellation of the wave with resultant drop in both volume and speaker efficiency. Treble-range speakers are known as **tweeters,** bass-range speakers as **woofers.**

Pickup Cartridges and Styli

All first-quality **pickup cartridges,** like the one shown in Figure 11.8, work on electromagnetic principles. Because of the very light loading of the stylus as it tracks the record groove, extremely little mechanical energy is available for operation of the pickup mechanism, which must be very light and easily moved.

Figure 11.7 Loudspeaker System The Pioneer Electronics S-510 system incorporates a heavy-duty 10 inch woofer for base frequencies, a smaller mid-range speaker of the permanent-magnet type, and a piezoelectric tweeter for high frequencies. The control on the model plate (upper right) equalizes their outputs. (Photo courtesy of Pioneer Electronics, USA, Inc.)

Two options are possible: Either the stylus can directly move a tiny coil in the field of a compact ceramic magnet, or it can move a tiny iron vane or something similar in the field of a powerful electromagnet. In the latter case the field is altered slightly, and this change causes an induced voltage to appear in its windings. The **stylus** is usually a diamond type with 0.5 mil tip radius. Its normal playing life is 600 to 800 hours.

Figure 11.8 Pickup Cartridge In this Heath Company stereo magnetic cartridge, the stylus, connector plugs, and pins are visible. (Photo courtesy of the Heath Company)

Microphones

The most widely sold quality microphone is the **dynamic microphone,** which resembles a miniature loudspeaker in reverse operation. Instead of a fiber cone, it has an aluminum or plastic diaphragm the size of a nickel or a dime. When sound waves strike the diaphragm, it and its attached voice coil move back and forth in the field of a small but powerful permanent magnet. The resulting induced voltage is then applied to a high-gain preamplifier in the main amplifier chassis to bring this very weak signal voltage up to a level that compares with those from tape recorders, AM/FM tuners, and phono pickups. A dynamic microphone exhibits ruggedness, excellent frequency response, and resistance to sound overload when used by the performer at close range.

AM/FM Tuners

The better AM/FM tuners use separate tuned circuits for the AM and FM sections. Separate coils are used for the mixer and high-frequency oscillator, and the switching takes place after frequency conversion has occurred. Separate detector circuitry is used for AM

The Pioneer Electronics model PL-S40 quartz-controlled direct-drive turntable provides not only exceptional speed stability but also automatic return of the playback arm to the starting position. (Photo courtesy of Pioneer Electronics, USA, Inc.)

and FM. The output of the AM detector is a monophonic signal. The output of the FM detector is usually a stereophonic signal unless monophonic FM is being broadcast. At least one stage of audio frequency amplification follows the detectors to bring up the signal level before applying it to the input jack of the main stereo amplifier. A tuning meter is supplied to ensure accurate tuning of both AM and FM signals. The meter is of particular value in tuning in a stereo signal on FM since incorrect dial tuning can cause serious audio distortion in the output.

In the latest AM/FM tuners, units share no common circuits, and many discrete components that were once visible have disappeared into one or two very large scale integrated circuits (VLSICs). The result is an apparent reduction in the complexity of modules like the multiplex decoder, essential in stereo FM reception. In the future, further use of large-scale integration will simplify design and reduce manufacturing costs.

Other Accessory Equipment

An important characteristic of a good record player is heavy-duty construction with a husky turntable and adequate bearings on which it can turn. The drive motor is especially critical. It should have at least two speeds, 33 1/3 rpm and 45 rpm, and maintain these speeds within very close limits. The slightest speed change during playback can cause a distressing pitch change most noticeable in classical music with its long, sustained tones.

Of critical importance in providing the best of audio quality and the longest possible life of the recording is the pickup arm, which securely holds the magnetic cartridge and also moves it across the record grooves. A well-designed pickup arm will have adjustable stylus weight on the record groove. It will have minimum friction in both its vertical and lateral bearings (pivots) and will maintain the pickup cartridge at a relatively constant angle in respect to the record grooves from the outside to the innermost groove. The best arms are equipped with silicone mechanical damping to prevent disc damage from accidental dropping of the arm and stylus on the playing surface of the record. This feature also makes it possible to play records that have become badly warped.

Room Acoustics

Frequently a comparatively inexpensive sound system will produce music of excellent quality in a certain room, while a very expensive sound installation may sound unbearable in another room. The variable factor is **acoustics,** the science of sound control.

The worst possible room in which to install a sound system would be one with no acoustic treatment—for example, a room with a plaster ceiling, a hardwood floor, and four plastered walls. Sound waves would simply bounce around in such an environment. The addition of a rug, some upholstered furniture, heavy drapes, and even some large potted plants makes a big difference. First, much of the energy of the higher-frequency sound waves is ab-

sorbed. Second, the wave fronts of other sound waves are broken up. The result is music that sounds more like the original in the control room or studio. Speaker placement can also be critical. Moving a speaker cabinet just a few inches can make a difference. If problems persist, an expert might recommend installation of a few well-placed acoustic tiles on walls or ceiling. Room acoustics can enhance or diminish the performance of an otherwise excellent sound installation.

DEVELOPMENTS IN RECORDING AND PLAYBACK

In the audio field a number of changes in basic technology and equipment, as well as in the entire recording-playback technique, have taken place within a few years. Rapid change is to be expected in any area of electronics where large sums are budgeted for research and development.

New Coatings for Tapes

For decades it had been accepted that iron oxide particles of varying degrees of fineness were the ideal coating for all magnetic tapes. Then someone tried using a trace of chromium oxide, and the high-frequency response improved dramatically. Cobalt seemed to offer promise, too, but was more expensive and required special recorder adjustment.

In 1978 the 3M Company, which had pioneered recording tape, announced an entirely new magnetic tape called Metafine. The tape's metallic particles are so fine that if a teaspoonful were tossed into the air, they would combine with the air's oxygen with explosive violence. Each particle must be given a coating of plastic before being incorporated into the tape's coating. Metafine is much more expensive than conventional iron oxide tapes but greatly increases the **dynamic range** (range between the softest and the loudest signal) that can be used in the record-playback process. Some recorders cannot record on Metafine-coated tape, but they can play back prerecorded demonstration tapes and show great improvement in audio quality.

Cassettes

After years of indecision, the recording industry is now favoring production and sale of **cassettes** and is rapidly retiring from further production of tape cartridges. Tape cartridges used for stereo frequently had as many as 8 parallel tracks, or channels, on a single 1/4 inch tape. Since the tracks were so close together, cross-talk from an adjacent channel was common. Because so little magnetized path was in contact with the pickup head, the playback tended to be noisy and of marginal quality. A modern, 4-track audio cassette (2 tracks in each direction) no longer has the cross-talk between channels, the need of highly critical head alignment, and the noise. Audio quality is greatly improved. The audio cassette approximates the quality of the finest LP recording.

Digital Technology

Digital technology is applied to the recording and reproduction of sound as follows: Sound waves are turned by a microphone into a pulsating DC current that varies at a voice or music frequency rate. This current is referred to as the **analog form** of the sound waves. Then, by suitable electronic circuitry, the current is turned into a properly timed sequence of DC pulses of electricity—the **digital form** of the sound waves. The sound is then amplified as a digital signal. Since the amplifying of a digital signal creates little noise or distortion and its frequency response is virtually unlimited, superb amplification takes place. The digital signal is now converted back to its analog form and can be used to operate a loudspeaker, a recording head in a recording studio, or a tape recorder. The process is expensive and complex, but where very high quality is essential, it is unequaled.

Within a few years major recording companies will use digital technology to make their master discs

and tapes, but smaller companies will continue to find conventional analog technology satisfactory. Because a comprehensive practical background in digital electronics and audio engineering, with special emphasis on disc and tape recording, is required, few technicians are working in the field of digital technology. It seems unlikely that their numbers will expand.

Compact Discs

The new **compact disc**, introduced in 1982 after years of research and development by Sony and Philips, is a round "sandwich" of clear plastic 4.7 inches in diameter (Figure 11.9). It encloses an extremely thin, metallic coating in which the program material has been permanently embedded by the burning action of a flickering laser beam used in the recording process. Disc speed averages 400 rpm, and one side of the 12 centimeter record provides about an hour of playing time. The record-playback process closely resembles the digital enhancement procedure NASA uses in transmitting camera information from space probes to Earth.

In the CD recording process, sound waves striking the microphone produce in its output an electrical wave that corresponds to the waveform of the sound wave and is called the analog. After amplification by active devices, the larger analog wave is subjected to a sampling, thousands of times per second, by a nearly infinite number of DC pulses of great accuracy. The result is a series of encoded digital "words" that are then used to modulate (vary) the intensity of a powerful laser beam brought to a pinpoint focus on the revolving disc. The laser's tremendous heat burns pits and rough spots in digital form into the encapsulated metallic layer that contains the program material. In an hour-long recording, about 6 billion bits of information are linearly encoded along the spiral path of the laser beam. The result is remarkably accurate reproduction on playback as well as great dynamic (volume) range without distortion.

The CD playback unit employs a lower-power laser as a pinpointed light source to read the pits, rough spots, and flats (areas relatively unchanged by the recording laser beam) of the disc's surface. The laser can be arranged to focus its beam through the disc or to direct the beam at an angle from above. The beam (direct or deflected) then strikes a special photovoltaic cell that serves as a transducer to change the varying light information into varying digital data. A special digital-to-analog IC recovers the original analog waveform, which can then be fed to the input circuit of any good stereo amplifier.

Figure 11.9 Compact Disc Jointly developed by Sony and North American Philips, the compact disc is dwarfed by a conventional 12 inch LP album, yet it offers longer playing time and improved frequency and volume range. Protected by clear plastic and read by a laser beam, it will never wear appreciably. (Photo courtesy of Sony Corporation)

The superiority of the CD over ordinary recordings is apparent on playback. Even very loud musical passages are essentially distortion free. The music has a new clarity and definition. Since the program data are sealed between two layers of plastic

and read by a laser beam, they are without background scratches, clicks, and pops.

Future of Digital Recording

The principal merchandising problem of the CD is the high cost of the disc (a little less than $20) and of the special playback units (a little less than $1000). A few record companies have started limited production of 12 inch LP discs in which the original (master) recording was made by the digital process used for compact discs, and from this digital master an analog stamper is produced. The vinyl discs produced by this hybrid process are better than conventionally recorded albums in terms of clarity and freedom from distortion. However, they cost four or five times as much as conventional albums and are subject to groove wear from the playback stylus.

In the future, premium LPs and albums will probably be digitally recorded. However, new VCRs have obtained remarkable improvement in audio quality by frequency-modulating the audio component along with the video and then demodulating the resultant carrier wave. This analog approach will surely also be tried with conventional records.

SUMMARY

The two basic technologies used for the recording and reproduction of sound—the rotating disc with program information carried in its undulating grooves and magnetizable material moving before an electromagnet—have been known in principle for nearly a century. Edison and Berliner had worked out many of the problems of disc recording and playback by 1890 and had given public demonstrations of early models of the phonograph. At about the same time, Poulsen was able to demonstrate the feasibility of magnetic recording and playback, although handicapped by the fact that his recording medium (a steel wire) cut into the pole pieces of his record-playback head.

The capture of two magnetic tape recorders from German noncombatants during World War II spurred interest in mass production of tape recorders. The earliest units resembled reel-to-reel tape machines: A paper tape coated with iron oxide passed from one reel, past a record-playback head, to the take-up reel. This type of machine was the prototype of all audio and video tape recorders.

A modern home listening center usually consists of three components: input devices (program sources), audio amplifier (stereo), and output devices (loudspeakers). The most common input device is a highly stable turntable equipped with a sensitive magnetic pickup. Also common is an AM/FM tuner that resembles separate AM and FM receivers sharing a common power supply. The stereo audio amplifier is made up of two identical amplifiers (left and right) of about six stages each, with a that frequency response from 20 to 20,000 hertz. Power output is in the range 15 to 100 watts per channel. The most common output device consists of two or three permanent-magnet loudspeakers of various diameters mounted in a heavy wooden cabinet. Loudspeakers are connected in parallel and provided with individual controls to adjust their volume levels.

The quality of disc recording was greatly improved in 1925 by the introduction of the all-electronic Western Electric process. Shortly after 1950 Dr. Peter Goldmark and CBS Laboratories introduced the all-vinyl long-playing record (LP), which increased disc playing time, extended frequency response, and reduced background noise. Until Sony introduced its digitally recorded compact disc (CD) in the early 1980s, the only other improvement in record albums and individual discs was the conversion of the record-playback process from monophonic to stereophonic. Digital discs are both recorded and played back by a varying laser beam. They offer amazingly low distortion, greatly expanded dynamic range, and no appreciable disc wear even after extended use. However, digital discs can be neither recorded nor played back on conventional analog audio equipment, and custom-produced digital audio equipment is still very expensive. Digital discs will probably dominate the record industry in the future.

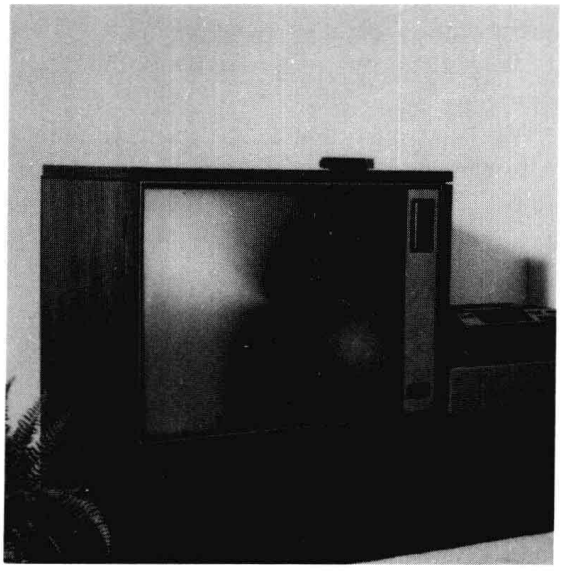
KEY WORDS AND CONCEPTS

acoustics
analog recording
cassette
compact disc
digital recording
distortion
dynamic microphone
dynamic range
equalization
erase head
high fidelity
hill-and-dale recording
lateral recording
loudspeaker
monophonic
phonograph pickup
pickup cartridge
push-pull circuit
recording head
sound waves
stereophonic
stylus
tweeter
volume unit (VU meter)
woofer

SELF-TEST QUESTIONS

1. Although Edison is credited with the invention of the phonograph, Berliner contributed at least three major improvements to the record-playback process. What were they?
2. What was the significance of the Western Electric recording process introduced by the Victor Talking Machine Company in 1925?
3. How did the improvement in the audio circuitry of AC-powered radio receivers make possible improved reproduction of phonograph records?
4. Describe what happens when a recording tape with its iron oxide coating moves over a tape deck's recording head with its constantly varying electromagnetic field.
5. Describe what happens when a recording tape is played back.
6. What is the difference between a monophonic and a stereophonic audio amplifier?
7. Define high fidelity.
8. The output transistors of a typical stereo amplifier develop a large amount of audio power at a comparatively high voltage and low current. What electrical transducer converts this power to the low-voltage, high-current AC power required by the loudspeakers?
9. Why do most modern sound systems use a number of loudspeakers of varying size rather than one or two large speaker units?
10. Why does mounting a loudspeaker in a cabinet improve its effectiveness in converting electrical waves into sound waves?
11. What is the average life expectancy of the diamond stylus mounted in a typical phono pickup cartridge?
12. Why is a turntable of constant speed essential in the playback of recorded music?
13. What are some simple steps that an audio hobbyist might take to improve room acoustics?
14. How is record wear reduced to nearly zero in the playback of digitally recorded CDs?

12 Video Systems and Electronics in Music



General Electric's Widescreen 1000 has a viewing screen over three times the size of the typical TV set sold today. Although it uses a standard GE television chassis, the 13 inch picture tube is designed for exceptional brilliance and clarity. The tube's image is magnified by a special high-definition lens system. (Photo courtesy of General Electric)

Video Tape Recording and Playback
- Review of Magnetic Recording
- Video Recording and Frequency Response
- Helical Scan Tape Recorders

Video Disc Recording and Playback
- A Review of Lasers
- Recording and Playback Process
- Playback Alternatives

Electronics in Music
- First Electronic Musical Instrument
- Electromagnetic Pickups
- Electronic Organs
- Electronic Pianos
- Music Synthesizers

Summary

Key Words and Concepts

Self-Test Questions

In the early 1950s television was rapidly expanding, but communications technology had not kept up with this expansion. Most of the programming centers were on the East Coast, and coaxial cable and microwave relays had not as yet closed the last remaining gap to the Pacific Coast. The nation's time zones added to the dilemma. The entertainment evening on TV was half over in Boston when San Franciscans were just sitting down to dinner. Some form of program delay was of maximum importance, and as a solution the networks turned to motion picture film. Film cameras loaded with black-and-white film were focused on a special control-room monitor tube of unusual brilliance known as a **kinescope.** The developed and printed film was called a *kinescope recording*, or simply *kine*. It was then ready to be air-expressed to West Coast TV network centers or to independent TV stations.

The visual quality of most kinescope recordings was very poor. Blacks were very black; whites were very white; intermediate shades of gray, in which much of the fine detail of a good monochrome TV picture is rendered, were lacking. Kinescope pictures were also fuzzy and grainy. The TV broadcast industry was willing to pay millions for electronic equipment that would record and play back visual materials with such clarity and fidelity that viewers would be unable to distinguish the recorded version from the original. Ampex, a small San Francisco firm, demonstrated the first successful videotape recorder and playback machine in 1956. Their previous experience in designing high-quality audiotape recorders gave them an advantage in designing a videotape recorder based on the same principles.

VIDEO TAPE RECORDING AND PLAYBACK

In this section we examine the recording and playback of visual information. We deal with the subject in terms of its analogy to audio recording and playback, with which we are already familiar.

Review of Magnetic Recording

When a cellulose tape coated with iron oxide is moved at a constant speed past an electromagnet that has an audio frequency current pulsating through its windings, the varying magnetic field rearranges the molecules of magnetic coating into a pattern corresponding to the voice or music being recorded. For playback, the tape is rewound and then passed over the same electromagnet at a constant speed. As it passes over the head of the electromagnet, this time connected to the unit's input circuit, the magnetic pattern on the tape induces a weak electrical current pulsating at a voice or music frequency. The current can then be amplified to full loudspeaker volume.

A frequency range of 20 to 20,000 hertz is easy to record and play back with modern audio equipment, especially with a tape speed of 7 1/2 inches per second (ips), the original standard of reel-to-reel machines. But visual materials require a much higher frequency limit—perhaps as high as 3.5 to 4.5 megahertz. In general, the higher this limit, the better the rendering of fine detail.

The most important factor in determining the high-frequency response of a recorder, excluding its amplifying equipment, is the unit's **writing speed**—that is, the relative speed between the recording head and the recording tape. In the system just described it would simply be 7 1/2 ips.

If it becomes necessary to expand widely the high-frequency response of a tape recorder, the writing speed must be greatly increased. Such expansion is not necessary or desirable at normal audio frequencies 20 to 20,000 hertz. However, at video frequencies, where the upper limit must be pushed as high as 3.5 to 4.5 megahertz, the writing speed must be tremendously increased—perhaps as high as 400 to 500 ips.

At first engineers tried to speed up the tape speed of conventionally designed audio recorders. The results were very disappointing. Even a 1000 foot roll of tape provided less than a minute's recording. Ampex engineers came up with the solution. Their ap-

proach, simple in theory but difficult to achieve technically, was to use a magnetic tape 2 inches wide moving from reel to reel at a speed of 15 ips—the writing speed of the audio recording heads used for the sound tracks only. The video recording head consisted of a revolving armature in contact with the moving tape and turning at thousands of rpm. It thus achieved the required high writing speed. The actual video writing speed was hundreds of inches per second. Video currents were led into the revolving armature by the use of brushes in contact with slip-rings as used in electrical machines. Playback was achieved by reversing the procedure and connecting the revolving heads to a video amplifier and monitor. Results were excellent, and viewers were no longer able to easily distinguish between a live and a recorded program.

This Ampex model AVR-3 quadruplex videotape recorder is typical of the video equipment used by the major TV networks for recording television programs in color. Note the color and waveform monitors in the control panel to the right. (Photo courtesy of Ampex Corporation)

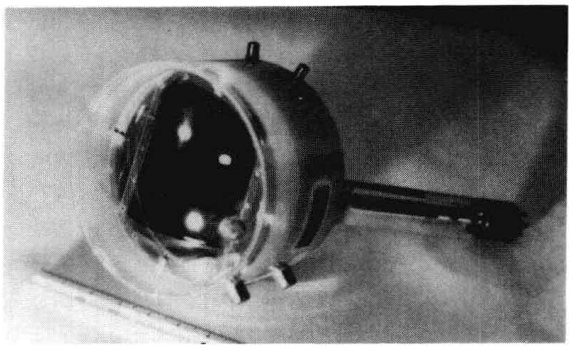
The first machines were monochrome (black and white), but after modifications in the mechanical design, excellent color recording and playback capability was achieved. In color use, absolute mechanical stability in the drive mechanism and the tape itself was a must. The slightest speed change could cause a noticeable change in the hue of a color.

Since the magnetic pattern was laid down by the quadrature (four-segment) recording head across the axis of the tape, it was called **transverse recording.** Large machines using this principle were expensive, which confined their use to the TV networks and the larger TV stations. Manufactured by Ampex and RCA, many of the machines are still in daily service. However, the high-speed rotating heads require constant and expensive maintenance, and they tend to dig into expensive videotape. Many are being replaced by simpler, high-performance videotape recorders (VTRs) of the helical scan type, which are less costly than the larger units. Unlike **quadrature machines,** which lay down their magnetic patterning at right angles to the axis of the tape, helical recorders make their recordings at an oblique angle to the tape's axis. In the mid-1980s quadrature machines are seldom used, and improved versions of the helical scan recorders are finding widespread acceptance.

Video Recording and Frequency Response

A VTR unit must have an extensive frequency response, commonly extending between 50 hertz and 3.5 to 4.5 megahertz. Recording what is seen by a television camera on magnetic tape is infinitely more complex than the recording of voice or music, yet the same basic principles are involved.

A studio scene is brought to a focus on the sensitized screen of the TV camera tube. Some areas are black, others white, still others intermediate shades of gray. Each area has a corresponding electrical charge that can be released by a scanning electron beam. The result is a pulsating DC current in the



The first commercially produced TV camera tube was the iconoscope. The mosaic, or sensitized area (dark rectangle), measures about 3 1/2 inches by 5 inches.

external circuit. After processing and video amplification, the current can be used to feed the revolving heads of the VTR. The varying magnetic field surrounding the heads then lays down a corresponding magnetic pattern on the videotape. On playback, the field surrounding the magnetic patterning on the tape induces a weak but corresponding pulsating current in the revolving head, now switched to the input of the instrument's video input circuit. After video amplification and processing, it can be used to feed the input of a TV monitor, a TV transmitter, or a network.

When a great deal of information is to be recorded in a very short length of time, the system must be capable of recording a wide range of frequencies. The amount of spectrum a certain system can handle is known as its **bandwidth.** Videotape recorders that can handle a range of frequencies from 50 hertz to 3.5 or 4.5 megahertz are capable of recording sharp, clear pictures with a large amount of fine detail, or **resolution.** Poor equipment with inferior bandwidth will invariably produce images with poor resolution.

Difficulties in obtaining sufficient bandwidth to produce adequate picture resolution held back until recently use of the two most promising approaches to video recording: the helical scan tape recorder and the video disc.

Helical Scan Tape Recorders

Helical scan recorders get their name from the helical (spiral) configuration that the magnetic tape takes as it moves from the supply reel, past the recording head, and to the take-up reel. Since the rapidly revolving recording head is enclosed in a slotted metal cylinder for protection and the tape must entirely surround this cylinder, the tape forms a helical shape as it travels past the head.

Unlike quadrature tape recorders, helical scan machines lay down a magnetic pattern at a much shallower angle. For that reason they are also sometimes known as **slant-track recorders.**

Compatibility and Videotape Standards

Unlike the manufacturers of audiotape recording equipment (who early standardized 1/4 inch tape moving at a speed of 7 1/2 ips), the helical scan VTR industry has been plagued by incompatibility, or conflicting standards of tape width and tape transport speed, as well as variance in the angle of the slanting pattern laid down by the recording head. For example, it is usually impossible to play a videotape recorded on the machine of one manufacturer on the machine of another manufacturer. Frequently tapes recorded on the helical scan machine of a certain manufacturer cannot be played back on an earlier or later model VTR by the same manufacturer. The industry is moving to near-compatibility, but full compatibility is still years away.

For years the description reel-to-reel (R/R) adequately described videotape equipment, and the acronym VTR was adequate to describe any videotape recorder. However, after manufacturers devised a system for mounting the supply and take-up reels in a plastic case, or **cassette,** for plug-in insertion, VTRs became known as VCRs (video cassette recorders).

While audio recorder formats are standardized,

a variety of video formats are in daily use both in TV broadcasting and in nonprofessional application. For the home entertainment market, most of the industry is adopting two basic formats: the Beta, developed by Sony and Zenith, and the VHS system, developed by Panasonic and RCA. These formats, however, are mutually incompatible. A Beta recording cannot be played back on a VHS machine, and vice versa.

Equipment for Home Use

Typical of the well-designed VCR equipment aimed at the home entertainment market is Sony's model SL-5100 Betamax machine (Figure 12.1). Originally offered in 1979, the Betamax has undergone many improvements that reflect advances in electronics.

A major change is increased flexibility in programmability. For example, model 5100 permits up to five hours of recording on a single L-830 Beta cassette. A digital clock/timer makes it possible to record the same show every day or every week. A form of low-frequency radio communication, its infrared remote control gives watchers instant access to all SL-5100 tape transport and special effects functions from anywhere in the room. High speed (nine times normal speed) permits rapid back-and-forth scanning for the quick selection of the exact video sequence desired. Full-logic microprocessor controls enable all switching and control functions to be made with fingertip pressure, a sensitive VHF-UHF tuner is built in to allow 14 button express tuning.

Accessories (Figure 12.2) include a color camera (a chassis input jack is provided) and an audio adapter to permit stereo sound, when FCC-authorized, to be channeled into the system. Visual quality of recorded material is so high that it is difficult to distinguish between recorded and directly viewed material. Audio quality is improved through a new

Figure 12.1 Sony Betamax VCR Typical of today's first-line helical scan machines, this VCR has a digital clock/timer for automatic shutoff and recording. Up to five hours of programs may be recorded on a single cassette. An infrared transmitter (foreground) provides instant remote control. (Photo courtesy of Sony Corporation)

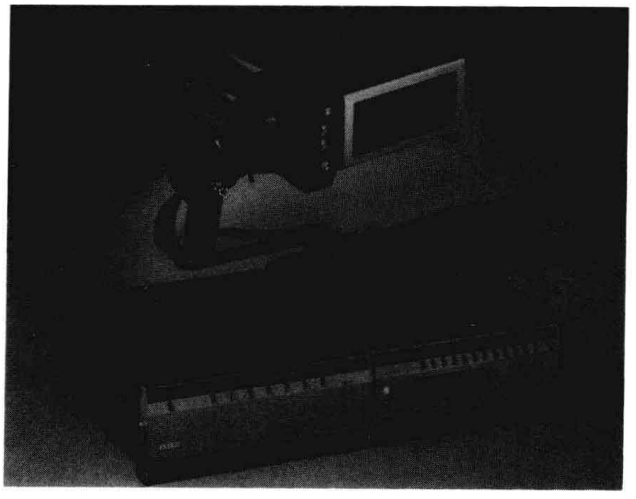


Figure 12.2 RCA's Convertible VCR with Camera Modularized for use at home and in the field, the RCA VFPI70 convertible VCR separates recording from playback functions for ready portability. At home, it comprises a full-feature VCR. The CC010 camera features a zoom lens with a choice of automatic or manual focusing. (Photo courtesy of RCA)

technique in which audio and video signals frequency-modulate an oscillator. The resultant radio frequency wave is used to record the magnetic tape. In playback, an FM demodulator unit recovers the audio and video information, which are separately processed.

Prerecorded Programs and Program Storage

Eventually Sony hopes to have a library of prerecorded tapes covering a variety of subjects. They will be available to Betamax owners on either a rental or a purchase plan. RCA is also investing large sums in a similar library. However, RCA seems to be making most of its acquisitions in terms of program material suitable for transfer to its 12 inch video discs, which will have up to an hour's playing time and retail for about $20 or more.

The advertising pitch behind the promotional campaigns of VTR manufacturers is the possibility of recording your favorite TV programs while you are at work and playing them back during your leisure hours. Such recording for personal use is not construed as a copyright violation. Compact color TV cameras are also manufactured for use with many VTRs, but often their cost equals, or may even exceed, that of the complete video recorder unit.

VCR Sales

During the recession of 1981–1983, VCRs continued to sell briskly in both Beta and VHS formats. The average sale price was about $700, although Sanyo offered a unit for only $329. By late 1983 many brands of VCRs were selling for about this price or even less. For many years Sony's Betamax dominated the marketplace, and its prerecorded tapes were on sale everywhere. By the mid-1980s Sony had lost its position of dominance, and nearly all VCR manufacturers were using the VHS (video home system) format, which offers longer playing time per reel and avoids payment of a royalty to Sony. VCRs are still not compatible in terms of tape width, angle of recording scan, and many other particulars, so tapes are not exchangeable among different machines.

In a 1984 landmark decision that should stimulate VCR sales, the Supreme Court ruled that videotaping a broadcast television program for personal use was not an infringement of the U.S. copyright law. Motion picture associations had contended that taping such programs was an infringement of copyright and had asked that royalties be paid. In a subsequent interview Motion Picture Association President Jack Valenti implied that the association would ask Congress to impose an excise tax on the sale of all blank videotape reels. Proceeds would be distributed to artists in lieu of royalties.

An RCA engineer examines a metallized disc that can store over 10,000 separate color images in a band only 1 inch wide. In a fraction of a second, each image can be located and projected on a screen. (Photo courtesy of RCA)

VIDEO DISC RECORDING AND PLAYBACK

The video disc, analogous to the LP record in the sound recording field, has been touted for a dozen years or more as the ideal medium for the recording and playback of visual information. It can take either of two approaches: the electromagnetic or the laser beam. Electromagnetic recording on disc is largely confined to instant replay of sports events on television. In the following sections we discuss methods of video recording that employ a laser beam to record the program and later to play it back.

A Review of Lasers

Basically, a laser is one of the many transducers we encounter in electronics that change one type of energy into another. When certain solids, liquids, and gases are subject to intense, pulsating electromagnetic fields such as could be produced by pulsing a heavy current through a coil, the energy level of the electrons that make up their atoms undergoes rapid changes. Under certain conditions, these changes give rise to the radiation of **coherent light waves**—that is, waves of a near-single wavelength all vibrating in the same plane. Coherent light waves produce an intense, highly concentrated beam of light that can be focused to a pinpoint. This intense laser beam can be readily modulated (varied in intensity) at an audio or video frequency rate and put to work in video recording on vinyl or acetate discs.

Recording and Playback Process

Figure 12.3 shows an optical video recording system. Figure 12.4 shows an optical video playback system. Although all the basic components are shown, some additional refinements need to be added to produce a practical system. The recording process is basically as follows: The pinpoint of light from the laser beam is first varied in intensity at a video frequency rate (in accordance with the areas of light

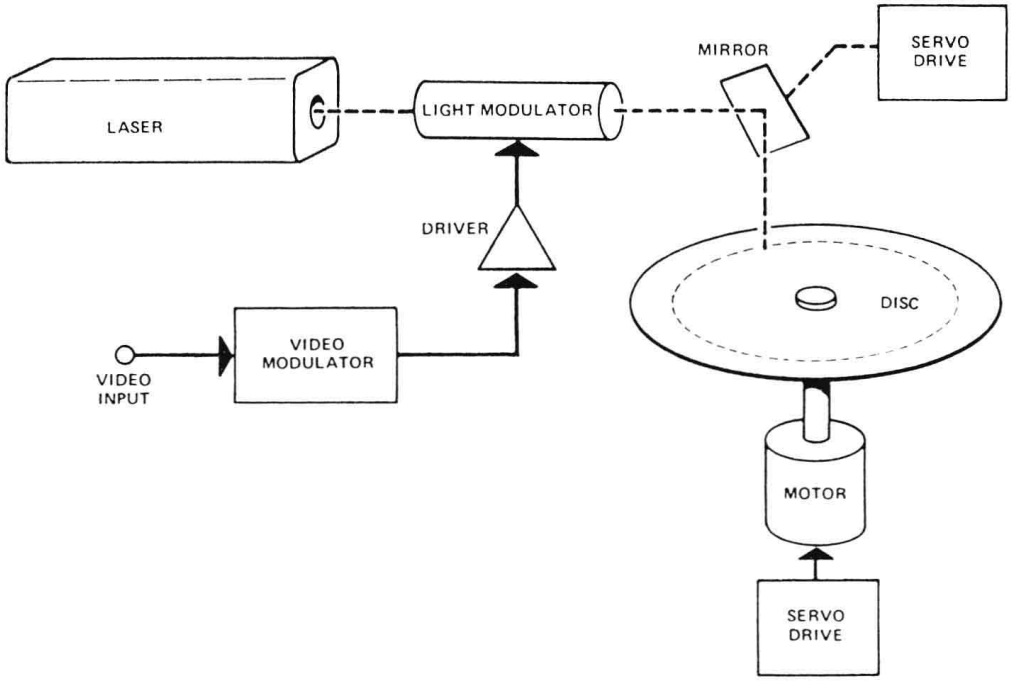


Figure 12.3 Optical Video Recording System This analog system uses a modulated laser beam to record a video disc.

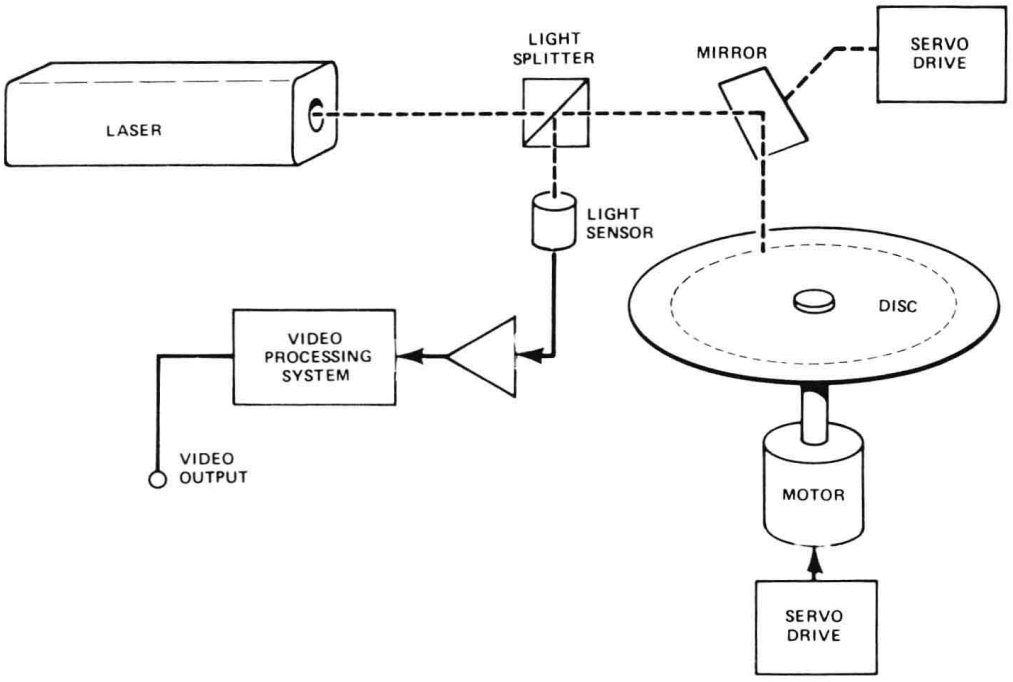


Figure I2.4 Optical Video Playback System This analog system uses a low-power laser to recover the video information from the disc.

and shadow appearing on the TV camera's sensitized screen). It is then reflected downward by a mirror, which steadily advances the laser spot on the disc in a continuing spiral pattern from the outside to the inside of the disc.

The master disc may be an absolutely flat piece of glass that has been coated with a very thin film of aluminum. As the laser beam advances along its spiral path, its intensity varies at a visual frequency rate, burning pits or blisters into the thin metallic film. These deformations contain the visual information as well as the audio data.

In playing back the master recording, the laser beam traces the same spiral path it followed during the recording process. A continuously varying amount of light is reflected back from the pits and blisters on the disc's aluminum-coated surface. This light is returned by the mirror to the **light splitter** (a double prism that splits a beam of light into two or more parts). The light splitter diverts the varying light beam to a supersensitive and fast-acting photoelectric cell. The output of this cell is a pulsating DC current varying at a visual frequency rate. A TV monitor converts this signal to a TV picture. The stereo audio program material is recovered in a similar way. It is also possible to modify the playback process by having the light source below the disc and directed toward the photoelectric cell.

The video-audio system we have described, although analog in its approach, almost equals the quality and frequency response of a much more complex and expensive digital system. Therefore, almost all 1984 VCRs use this system for their audio channel as well as for their video. The result is a vast improvement in audio frequency response and dynamic range.

Playback Alternatives

If the disc were to be pregrooved clear mylar or similar plastic material, changes in the groove material caused by the heating effect of the laser beam

could be used in a different way in the playback process. A photoelectric cell could be mounted under the rotating disc. The amount of light it received from the laser beam, shining through the record groove would vary at a video frequency rate. The cell would then turn the light back into pulsating DC varying at a video frequency rate.

It is also possible to use a metal stylus in a conventional grooved record that contains visual information in the record grooves. In RCA's SelectaVision disc playback machine (Figure 12.5), the metal-tipped stylus is part of an electrical circuit that is completed by the varying metal coating in the bottom of the record groove. Disc rotation produces the effect of a variable capacitance whose capacity varies at a visual frequency rate. If the varying capacitance is placed in parallel with a radio frequency oscillator circuit employing a transistor, this unit is frequency modulated at a visual rate. It becomes a simple matter, then, to use a basic FM detector circuit to convert this FM signal back into a pulsating DC current varying at a visual information rate. Applied to the input of a TV monitor, this current causes a monochrome or a color picture to appear on its screen.

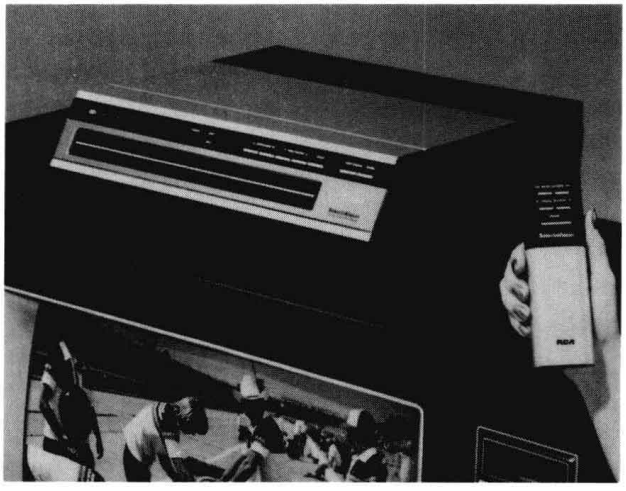


Figure 12.5 RCA SelectaVision Video Disc Player This unit operates on the principle of variable capacitance between its diamond stylus and the undulations in the record grooves. The record grooves frequency-modulate an oscillator, are amplified, and are then demodulated to provide excellent video and stereo audio. A hand-held infrared transmitter provides remote control. (Photo courtesy of RCA)

To achieve the high writing speed required for a disc frequency response that must extend as high as 3.5 or 4.5 megahertz, RCA's disc must turn at a speed of 450 rpm. Even with this high disc speed, up to 30 minutes of playing time can be obtained from each side of the record. Excellent monochrome and color quality is obtained. However, the high rotational speed may require that the disc be lubricated to prevent excessive stylus wear. Even then, about 200 hours of playback time is all that can be expected, and a new stylus costs about $70. As is true with all of today's video disc machines, only a playback facility is provided. The actual recording of a disc requires factory mastering equipment. Magnetic tape machines that provide both video recording and playback capability would thus seem to have an advantage, but many people are principally interested in prerecorded discs of professional quality.

All of RCA's 1983 video equipment, including both video disc and helical scan tape machines, carries the generic name SelectaVision. Therefore, it is necessary to be specific about which type of equipment is being discussed. In 1984 typical prerecorded video discs were selling for approximately $20, and the number of available titles was increasing rapidly. A few video games are also available.

ELECTRONICS IN MUSIC

Much of the basic technology applied to the design and construction of wide-range and low-distortion recording and playback equipment for use with both aural and visual program materials has been borrowed in part by the designers of electronic organs and pianos. The powerful audio amplifiers that such

instruments use to convert electrical waveforms back into sound waves through the medium of the loudspeaker are precisely the same as the units we discussed in the previous chapter on audio systems.

Electronics in music is a multimillion dollar business that continues to grow rapidly. It is consumer oriented and caters largely to those to whom high-quality reproduction of music produced by some specific intrument is of maximum importance. Cost may be a comparatively minor consideration. A small, but important market is also developing for specialized and sophisticated units like the Moog and ARP music synthesizers, principally used in larger recording studios and by groups interested in electronically created music.

Many new jobs will continue to develop in the installation, tuning, adjustment, and repair of musical electronic equipment. Technicians with a good knowledge of music fundamentals and of modern electronics should readily find employment in this field.

First Electronic Musical Instrument

The first purely electronic musical instrument was known as the *Theremin*, named after its inventor and developer, Russian-born scientist Leon Theremin. Essentially it was an audio frequency oscillator followed by several stages of vacuum-tube amplification and one or more loudspeakers. Capable of producing only one musical tone at a time, the pitch of the tone developed could be varied over several octaves by the performer moving his or her right hand closer to or further from a vertical metal rod while the left hand controlled the volume. Like its mechanical predecessors, the slide whistle and the so-called musical saw, the Theremin had no fixed intervals of pitch such as would be provided by a keyboard instrument. The pitch of the tone produced depended entirely on the skill of the operator. A number of musical compositions were written for the Theremin and publicly performed, but interest lagged and today this instrument is only a museum curiosity.

Electromagnetic Pickups

The use of electromagnetic pickups to amplify the sound of steel-stringed instruments was the next application of electronics to music, and one that continues at a lively pace today, especially in the case of the guitar (Figure 12.6). Much of the technology had been worked out in the development of the electromagnetic phono pickup, and the basic principles needed little modification. An electromagnet was mechanically mounted on the instrument in close proximity to the vibrating steel strings. As the strings moved mechanically, they caused slight changes in the configuration of the field surrounding the permanent magnet. The changes were equivalent to movement of the magnetic field, and a pulsating DC voltage was induced in the winding wound over the magnet and carried through a shielded cable to a high-gain audio amplifier. Later designs provided a single electromagnetic pickup for each string of the instrument and elaborate shielding to prevent hum

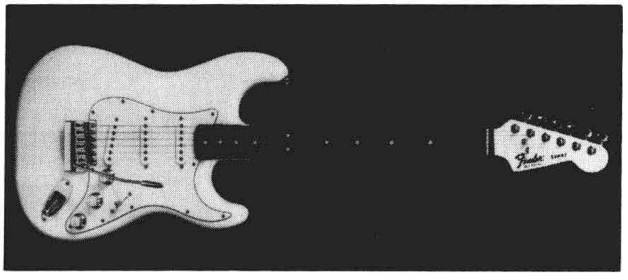


Figure 12.6 Electronic Guitar Three electromagnetic pickup units are mounted immediately under the strings of a Fender Strat 6-string guitar. Knobs at lower right control tone and volume. (Photo courtesy of Rhodes Keyboard Instruments)

pickup. Today, many of the best-quality units have built-in integrated circuit chips in the pickup itself, permitting the use of a very long cable between instrument and amplifier without picking up stray noise or hum.

Electronic Organs

The first complete electronic musical instrument was the Hammond organ, designed by San Francisco inventor-engineer John Hays Hammond. First marketed about 1935, it was based on the principle of the alternator, or AC generator. A synchronous electrical motor drove a shaft that was mounted behind each keyboard. Behind each key and firmly mounted on the shaft was what appeared to be a toothed iron gear that rotated in the field of an electromagnet for each individual key. Each toothed gear for a certain note on the keyboard differed slightly from its neighbor a tone higher or lower. As the toothed iron wheel revolved in the electromagnet's field, its serrations caused variations in the magnetic field that caused an AC current to flow through the electromagnet's windings and external circuit. This current is essentially a pure sine wave having little harmonic development. When amplified, it sounds like an orchestral flute, and it can be heard on all 44 keys of each manual (keyboard) and on the 11 keys of the pedal board. Orchestral instruments can be imitated quite closely by feeding a part of the output of the tone generators through tone filters made up of resistors, capacitors, and inductors connected in various patterns.

In recent times the trend has been away from the motor-driven alternator design and toward the use of master audio oscillators that provide more design flexibility, particularly in the development of new and varied tonal colors. Instead of the original Hammond alternator concept, all current organ manufacturers use a very stable master oscillator—an active device that operates on an audio frequency higher than the highest note on the manual keyboard. By the use of a series of specially designed ICs, this high audio frequency is repeatedly **divided down** to provide output on the precise frequency of the traditional 12-note scales. Larger Hammond organs have two 61-note manual keyboards and 25 notes in the pedal clavier. Spinet models have two manuals with nearly four octaves each and a pedal clavier with one octave.

The Hammond B-3000 electronic organ represents a great advance over earlier models, but one feature has been retained: the harmonic tonebars that provide an infinite degree of tonal variation. The instrument has two manual keyboards of 5 octaves each and a pedal clavier of 25 notes. (Photo courtesy of Hammond Organ Company)

Standard pitch throughout the Western world is the A above middle C on a standard keyboard instrument. This A has a frequency of 440 hertz. The A one octave above is exactly twice this frequency, or 880 hertz. The A one octave below is precisely half, or 220 hertz. This same relationship extends throughout the musical scale.

Harmonics

The fundamental output of each individual master oscillator is a **sawtooth wave,** so called because of its appearance when viewed on an oscilloscope. It is rich in harmonics (other waveforms that are two, three, four, or even five times the fundamental frequency). The presence of harmonics added to the fundamental frequency in varying ratios gives musical instruments their characteristic sound or timbre. For example, the flute tone is primarily a sine wave with little harmonic development. A violin is characterized by a strong third harmonic. The varying ratio of fundamental to harmonics also accounts for the tonal characteristics of brasses and reeds.

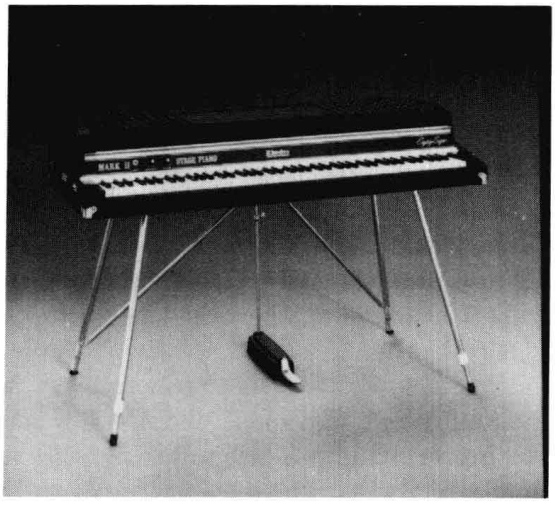
Organ manufacturers take the fundamental sawtooth waveform, which sounds somewhat like an oboe when amplified, and pass it through **formant filters** that strengthen some frequencies and weaken others. For example, a formant filter consisting of resistance, capacitance, and inductance in a proper configuration can be used to roll off the higher harmonics and leave a flutelike tone. Various harmonics can also be strengthened to produce a string tone. The variety of possible combinations seems almost endless.

Other Tonal Variations

Most instruments also offer various percussive effects, such as reasonable simulations of piano and guitar. The effects are obtained electronically by modifying the original waveform so that it is turned on very abruptly and then allowed to decay slowly. New integrated circuits also make other effects available. Some permit momentary storage of what has just been played in digital form and then meter it back in analog form at a predetermined rate. If metered back at a steady rate, it provides a wide range of reverberation, or echo. If metered back at an irregular rate, the music can acquire rhythms characteristic of the tango, the rumba, or even the waltz, all determined by the particuiar IC chip switched into the circuit. With the development of new, inexpensive ICs for these and related applications, the electronic organ will become easier to play, will have a wider variety of tonal color, and will be equipped with a wider spectrum of special effects.

Electronic Pianos

The electronic piano has much in common with the clavichord and the pianoforte of Bach's time. The keyboard and action are the same (wood is extensively used), the mechanical linkage is the same, and even the system of felts used in damping is much the same. The basic difference is that the felt-tipped hammers strike small strips of spring steel instead of strings supported on a metal frame. Each steel strip is solidly held at one end. The other end is free to vibrate mechanically when struck by the hammer. Typically, one of these metal reeds is about 4 millimeters in length and 1 millimeter wide at the point of attachment, and about 0.5 millimeter wide for the narrow portion that is actually free to vibrate. The tip carries a bead of lead-tin alloy that when filed,



This all-electronic Rhodes Mark II stage piano operates on the same basic principle as that of a condenser microphone. It requires an external power amplifier and speaker. Foot pedal controls volume. Bass-boost control is also provided. (Photo courtesy of Rhodes Keyboard Instruments)

permits precise tuning of the reed. Low-frequency reeds are larger physically. High-frequency reeds are proportionally smaller.

All reeds vibrate in very narrow slots cut into a long metal baseplate that extends the full length of the keyboard. A 60 volt DC potential appears between the reeds and the baseplate. As the reeds vibrate under the impact of the striking hammers, they form variable capacitors that charge and discharge at a rate that is a function of their mechanical dimensions. The very small pulsating DC voltage picked up off the steel reed as it vibrates is an electrical analog (replica) of the mechanical sound it generates in the air surrounding the reed, although this sound is very weak.

After amplification in a high-gain (but entirely conventional) audio amplifier, a loudspeaker turns this audio frequency energy back to sound waves. Although the key action is excellent and tonal quality pleasing, the electronic piano is not a true replica of conventional spinet piano sound.

The Moog Company's Minimoog music synthesizer uses rotary selector switches to replace unsightly patch cords and plugs. The five basic sections of the front panel are (left to right) controllers, oscillator block, mixer, modifiers, and output.

Music Synthesizers

Although RCA pioneered the so-called **music synthesizer** as far back as the early 1950s, it remained for R.A. Moog, Inc., to make it a single unit that could be replicated. The basic equipment required is as follows: audio oscillators capable of generating several types of waveforms, voltage-controlled audio oscillators, white-noise generators, variable-speed tape recorders, audio filters of varying characteristics, and one or more keyboards and control panels.

Musicians do not perform at the keyboard of a music synthesizer. Instead, they assemble bits and fragments of musical phrases from various sources and record them on tape. Then, a master recording is made as a composite of all the short segments. One master tape that became a true masterpiece was the recording of "Switched on Bach," completely a product of synthesis and now a collector's item.

New uses are still being found for music synthesizers. For example, they are highly effective in enhancing motion picture sound tracks and radio and TV commercials. The record companies that turn out millions of rock records each year place great reliance on music synthesizers for the unique quality they can add to rock recordings.

SUMMARY

The recording of visual information (video) on magnetic tape is remarkably similar in principle to the recording of voice and music (audio) on the same type of tape. However, video recording is a more complex and sophisticated process because a large amount of information must be recorded in a very short period of time. An enormous amount of data obtained from the TV camera represents in electrical form the studio scene being televised. Controlling pulses govern such functions as blanking and synchronization, and one or more audio channels are needed for audio information.

Increasing the amount of information a certain piece of equipment must handle requires a corresponding increase in the system's bandwidth (the total spectrum it can handle effectively). While a

bandwidth of 20 to 20,000 hertz is completely satisfactory for audio, the very minimum for a videotape recorder (VTR) is approximately 50 hertz to 4.5 megahertz for good picture quality. This bandwidth can only be obtained by a very large increase in the tape writing speed—the relative motion between the tape and the recording head. A writing speed of 7 1/2 inches per second is no longer satisfactory. Instead, a writing speed of hundreds of inches per second is necessary and can be obtained by rotating the recording head at a high rate of speed as the magnetic tape moves past it at 7 1/2 ips or some other modest speed.

The first successful VTR of professional quality (the Ampex model 1000) was so expensive that its use was confined to the TV networks and major TV stations. It laid down its magnetic pattern at right angles to the direction of the moving tape. A later development, the slant-track recorder, using less sophisticated electronics and mechanical construction, laid down its magnetic pattern at less than a right angle to the axis of the tape. Twenty years later it is greatly improved and is the basis of all low- and medium-priced VTRs on the market today. It is also known today as a helical scan VTR—a term that describes the path taken by the tape as it moves through the machine. The acronym VCR refers to a VTR whose magnetic tape is enclosed in a cassette.

A video disc looks something like a conventional phonograph record. However, it has no grooves but has a spiral path, burned by a laser, that contains program information. This path contains picture information, controlling pulses, and encoded voice and music data. The path is read by a needle-sharp laser beam and its associated optical system. To obtain the high writing speed required for good video response, the disc must turn at a speed of hundreds of rpm. Video discs can use either a laser beam or a stylus. At present both types of video disc players can reproduce only prerecorded discs. They are not equipped for making home recordings. Although video disc players that use a stylus rather than a laser beam are less expensive, both types are capable of producing full-color pictures of excellent quality. Audio quality is also excellent.

Electronic musical instruments fall into two broad categories. First are the vibrating metal strings or reeds that create changes in a magnetic field. These changes, in turn, produce a change in voltage. Second are the active devices that generate a basic tone that is altered and divided down by suitable electronic components. Electric guitars and electronic pianos typify the first category; electronic organs and electronic music synthesizers typify the second. In both cases the changing voltage obtained is fed to the input of a conventional audio amplifier, and suitable loudspeakers convert the amplifier's output to sound waves.

In a typical electronic organ an IC generates a sawtooth wave that is rich in harmonics and has a fundamental frequency just higher than the highest note on the manual. This note is divided down by other ICs to provide the needed 12 tones for each manual octave and for the pedal clavier. Formant filters made up of electronic components filter the harmonic-rich sawtooth wave to produce sounds characteristic of flutes, oboes, clarinets, and orchestra bells.

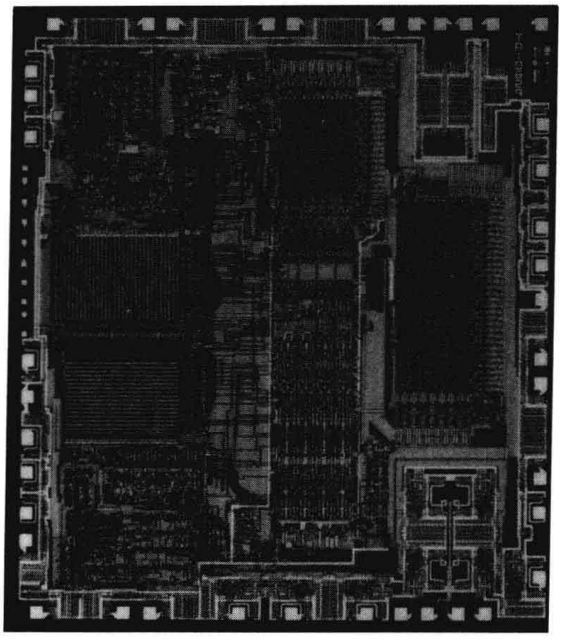
KEY WORDS AND CONCEPTS

bandwidth
cassette
coherent light
dividing down
formant filter
fundamental tone
helical scan recorder
kinescope
laser
light splitter
music synthesizer
quadrature machine
resolution
sawtooth wave
slant-track recorder
transverse recording
writing speed

SELF-TEST QUESTIONS

1. What is meant by writing speed in magnetic tape recording?
2. How did Ampex engineers achieve the high writing speed that made the first videotape recorders possible?
a. Why were the early quadrature types of videotape recorders seldom sold to smaller TV stations, to public schools, or to individuals?
4. Why are the videotape recorders produced for the home market known as helical scan machines?
5. What is meant by compatibility when the term is used to describe the tapes recorded on various types of videotape recorders?
6. Why is lack of compatibility a continuing problem with helical scan recorders?
7. What advantage do conventional VTRs using magnetic tape have over video disc playback machines?
8. Why are electromagnetic pickup units confined to use on musical instruments with steel strings?
9. Explain what is meant by this statement: The first Hammond organ was technically an electric rather than an electronic organ.
10. How many master oscillators are usually required in electronic organs to provide all the notes for the keyboards and pedals?
11. Why is an audio oscillator that provides a sawtooth wave, rather than a sine wave, preferred by most organ designers?
12. In organ design, what important function is performed by the formant filters?
13. Electronic pianos resemble acoustic pianos in construction. In what ways do they differ?
14. What are some of the new applications for music synthesizers in motion pictures, radio, and television?

13 Electronic Computers, Calculators, and Clocks



Used by many computer manufacturers, this complete central processing unit, or CPU, is mounted on a single silicon chip. (Photo courtesy of Intel Corporation, copyright 1977, 1975)

Historical Perspective on Computers
- European Advances
- Babbage's Analytic Engine
- Development during World War II
- Introduction of Transistors
- Microcomputers
- Categories of Computers

Digital Computers
- Basics of Computer Design
- Building Blocks of the Computer
- Home Computer Installations

Microprocessors

Electronic Calculators
- Construction of Electronic Calculators
- Trends in Visual Displays
- Trends in Calculators

Clocks and Watches
- Concept of Timekeeping
- Construction of Electronic Clocks
- Electronic Watches

Summary

Key Words and Concepts

Self-Test Questions

People have always sought ways to speed mathematical calculations and to minimize the drudgery they entail. Devices such as counting stones (calculi), from which we derive the modern word *calculate*, have been in use tens of thousands of years. The practice of counting on fingers and toes (digiti), from which we derive the modern word *digital*, extends even further back in time.

HISTORICAL PERSPECTIVE ON COMPUTERS

Well before the Christian era, the first true calculating machine was developed in the Far East, probably in China. Called the **abacus,** it was a simple device that enclosed in a wooden frame sets of colored beads that could be moved about on taut strings (Figure 13.1). It was based on what we call today base five math. By 460 B.C., use of the abacus had spread as far west as Egypt. Within a few centuries it was being employed by shopkeepers and merchants throughout much of the Roman Empire for routine business calculations. Even today it is widely used in Oriental countries. A skilled operator can perform addition and subtraction with amazing speed and accuracy.

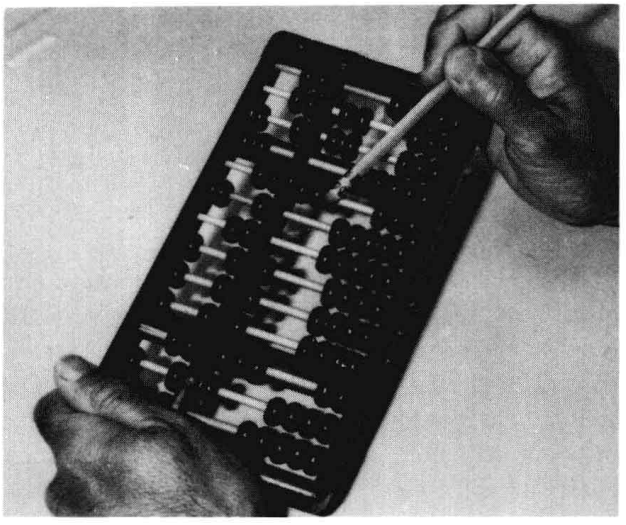


Figure 13.1 Abacus On the first true calculating machine, a skilled operator could perform addition and subtraction at a high rate of speed.

European Advances

After about A.D. 1100, Europeans began to draw heavily on the mathematical skills and number tables of the Arab world, which had been partly derived from the Greeks and the Indians. From the latter, Europeans acquired the concepts of zero and the decimal number system. Both sources contributed to European knowledge of algebra, geometry, and trigonometry. Their interest stimulated, the Europeans went on to make discoveries of their own.

Some of the greatest advances in mathematics in Europe immediately followed the Renaissance, when rapid developments in navigation, astronomy, and commerce required corresponding advances in technology and its applications. For example, Napier published the first logarithm tables in 1614. About 1675 Newton announced his calculus, which offered a new tool for solving complex problems that the physical sciences, particularly astronomy, were encountering.

Babbage's Analytic Engine

In 1833 Charles Babbage, a prolific British inventor, focused his attention on producing a piece of machinery capable of solving complicated mathematical problems quickly, without weeks of human concentration and computation. In principle Babbage's machine functioned much like a modern wattmeter: A series of meshing gears with varying numbers of teeth drove a large number of indicating hands or dials. It was far more complex than the wattmeter, however, and machinists of the era lacked the technical competence to turn out a piece of equipment that demanded such precision. Although Babbage's analytical machine was never completed, the working drawings show some remarkable insights into how a computer should be constructed and how it

might be used. Babbage's concepts are so well reasoned that they are still, in part, applicable to modern computers.

Development during World War II

During World War II rapid military expansion into many new fields of technology, particularly that of long-range missiles, produced computational problems that only some form of computing machine could solve reasonably quickly. IBM completed the installation of its first electromechanical digital computer, the Mark I, at Harvard University in 1944. One year earlier a research team at the University of Pennsylvania had started work on the first computer to use vacuum tubes rather than electromagnetic relays. Known by the acronym ENIAC (electronic numerical integrator and computer), it used base two math in its design and was initially used by scientists. However, as its application was extended to other fields, a serious design flaw became apparent. Since ENIAC lacked stored-in-memory instructions for basic operations, it was necessary to partially rewire sections of the computer for every new application. Since that time, it has been standard design practice to incorporate in a computer's memory certain basic instructions and number tables that are permanently retained. ENIAC did establish the advantages of using active devices such as vacuum tubes in place of the much slower magnetic relays.

Introduction of Transistors

By 1960 transistors were rapidly taking the place of vacuum tubes as control elements in computers. In the early 1960s new photochemical advances made possible the integrated circuit, which incorporated many transistors, their related circuit components, and interconnecting wiring on one tiny silicon chip. The age of miniaturization was arriving. New computers were constructed of amazingly compact modules. Another advantage of transistors and integrated circuits was that heat was no longer a problem. (Vacuum tubes had required the use of elaborate air conditioning equipment.)

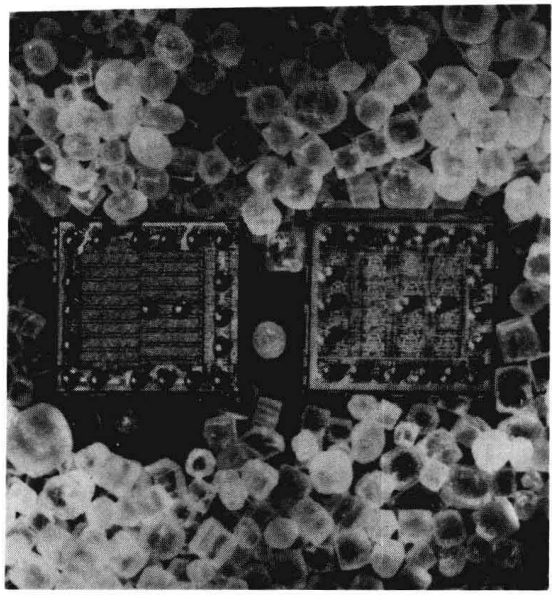
By the mid-1960s IBM's model 360 digital computer, requiring the equivalent of several office rooms for the basic unit and accessory equipment, was being widely employed by industry. Like modern computers, it was used for such applications as data processing, inventory control, and research. The IBM 360 became the model on which other manufacturers drew in developing their own machines. IBM was soon joined in the computer field by General Electric, Control Data Corporation, and Xerox.

Microcomputers

One final important breakthrough was development in 1970 of the technology to construct on a single silicon chip, measuring less than 1/6 inch square, all the components and circuits required in the arithmetic, logic, and control units of a digital computer. The three units make up what is known as the **central processing unit** (CPU) of the machine. The engineer who developed one-chip technology was Ted Hoff. His firm was Intel, a relatively obscure electronics plant in San Francisco's Silicon Valley.

Hoff's new miniature CPU, when linked with other units to provide memory, input, and output functions, resulted in the world's smallest digital computer. It could perform most of the functions of much larger computers but operated at a somewhat slower speed and lacked the capacity to handle the amount of data larger machines could readily process and store. For many applications, though, it was completely adequate.

Once the master design of the CPU unit had been perfected, copies were turned out by tens of thousands. As production increased, prices for the units plummeted. By the close of 1978, the price for a complete microcomputer, including video monitor, electronic keyboard, and cassette-type magnetic tape recorder, was comparable to that of a color TV receiver. In 1983 the same equipment could be purchased for 25% less and offered improved and more flexible operation.



Tiny electronic circuit chips surrounded by crystals of ordinary table salt have been a key factor in making computers faster and more economical. The logic chip (left) and the memory chip (right) are less than 1/8 inch square and are used in IBM's 3033 processor unit. (Photo courtesy of International Business Machines Corporation)

Categories of Computers

All computers fall into two general categories: **analog** and **digital.** Analog computers deal effectively with two or more varying quantities—for example, electrical power and motion or temperature and pressure. Analog computers were used by the U.S. military in the early 1940s primarily for gun control applications, where projectile speed and trajectory had to be matched to the speed and altitude of an enemy plane. Later models, employing vacuum tubes, were surprisingly rapid and accurate in their operation.

Digital computers are designed for accurate, high-speed manipulations with discrete numbers. At first they were slower than analog machines, but today they surpass analog performance.

DIGITAL COMPUTERS

Digital computers are electronic devices designed to work with discrete numbers and to perform at incredible speed mathematical functions such as addition, subtraction, multiplication, division, and the extraction of roots. They are also capable of comparing and rearranging data as well as storing the data for future reference. These operations are performed at blinding speed: 1/100,000,000 second or less on a full-sized computer.

A **minicomputer** is a scaled-down version of a standard computer. A **microcomputer** is a very small computer engineered around a central processing unit constructed on a single silicon chip and provided with minimal terminal equipment. Low-cost microcomputers are being sold by the millions as personal computers (PCs).

Basics of Computer Design

In early design work in digital computing, the fundamental concept was that a particular switch (either mechanical or electronic) would represent a particular number (digit) if the switch was in the *on* position. If the same switch was off, the value it represented did not exist.

For example, visualize a horizontal row of switches. Starting at one end of the row, let each successive switch represent a number twice as large as the previous one. In this manner we can use various combinations of switches to represent any number. That is, the number is represented by the sum of the values of the switches that are in the *on* position. The sequence of values for the first five switches is then 1, 2, 4, 8, and 16. Because the value specified for each switch of our series is an exact value, representation is very accurate. Therefore, digital computation, if properly done, yields highly accurate results.

Since each switch can only have two conditions (on and off), the combination can be represented by using 1 for the *on* position and 0 for the *off* position. The rules for performing addition, subtraction, and

like operations using only the numbers 1 and 0 are those of **base two,** or **binary, math.**

Building Blocks of the Computer

Figure 13.2, a block diagram of a typical computer system, shows that a computer has only three **modules,** or units. The most complex is the central processing unit (CPU), which today is supplied in ultracompact form as a single large-scale integrated (LSI) circuit. Next is a memory unit for storing data and instructions. Finally, input and output (I/O) circuits feed instructions and data into the machine and deliver processed data to the various pieces of equipment connected to the output terminals. Within all three major blocks are the subdivisions shown in Figure 13.3.

Central Processing Unit (CPU)

The CPU contains all the necessary electronic circuitry on a single silicon chip to carry out extremely rapid arithmetic operations with 4, 8, or 16 digit numbers, depending on the degree of complexity of the computer. Also, the computer has the ability to compare or rearrange the numbers at lightning speed and to store them in memory for future use. The section of the CPU in which these operations take place is known as the **arithmetic logic unit** (ALU).

Early digital computers used electromechanical switches (relays) to represent numbers within the computer. Many switches were required, and they took up much valuable space. IC technology has made it possible to provide on a single silicon chip vast numbers of specialized transistors, known as **gates,** that perform the function of switches. Partly because of the reduced size of the components, the speed of digital computation has increased rapidly. A few years ago large computers that performed 1 million operations per second were considered fast. Today some microcomputers can operate at eight times this speed.

Computer Functions and Addition. Unlike the approach used by a mathematician, the digital

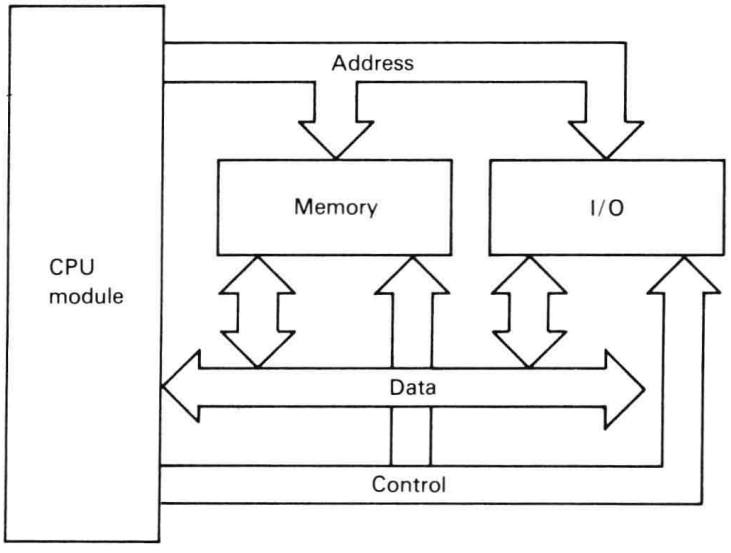


Figure 13.2 Block Diagram of a Typical Computer System All computers have only three basic modules: the central processing unit, the memory unit, and the input and output circuits. (Photo courtesy of Intel Corporation, copyright 1977, 1975)

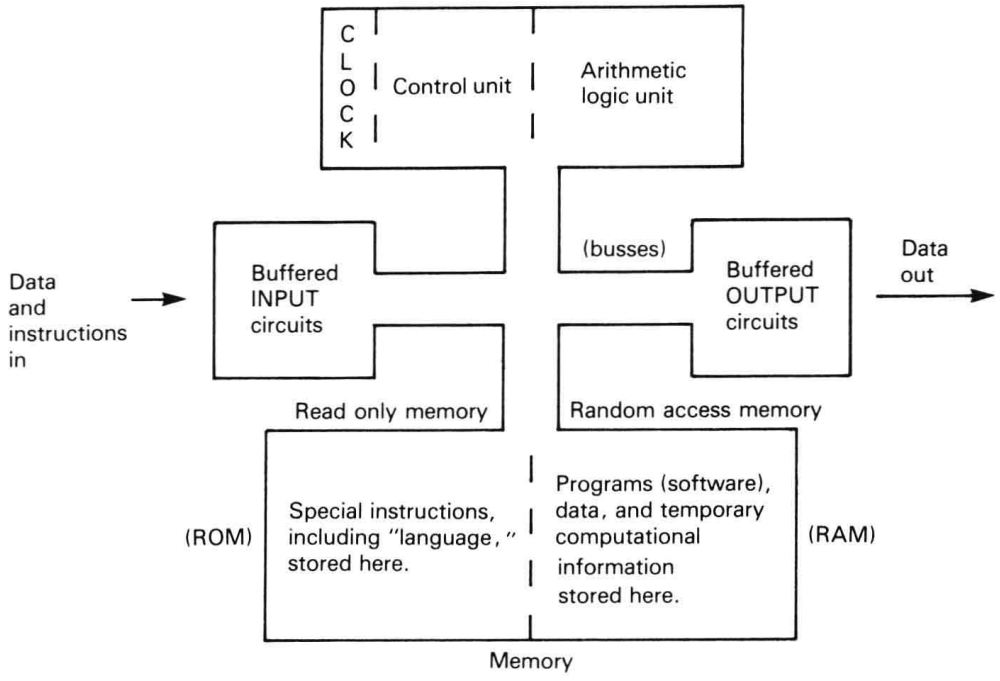


Figure 13.3 Block Diagram of a Computer System's Subdivisions The CPU includes the arithmetic logic unit and a precision timing unit that is built around a clock. Computer memory is basically divided into permanent read only memory and temporary random access memory.

computer performs all computations by one basic operation: addition. For example, the sum of 750 and 250 is obtained by conventional addition. To subtract 250 from 750, the computer changes the sign of one quantity and again adds. Multiplication is really rapid sequential addition. Division is rapid sequential subtraction (although mathematical shortcuts may be used instead). Many other mathematical operations can be performed by combining the addition, subtraction, multiplication, and division functions, but the CPU of the computer is only performing addition. Simple addition and subtraction take far less computer time than multiplication and division, which require a sequence of operations. Yet even these operations are performed very quickly.

Timing Unit. The CPU module contains a recision timing unit built around a clock. The unit is basically a conventional crystal oscillator circuit similar to those described in Chapters 9 and 10. The entire module is simply designated as *control*. Its primary function is to time precisely all steps spelled out in the program and to see that they are performed in the right sequence. In general, the higher the fundamental frequency of the quartz crystal, the more rapidly operations can be performed. Full-size computers can perform in 100/1,000,000 second (100 microseconds) or less. With minicomputers and microcomputers the clock timing is slower, but the amount of time required is still extremely small: 1/1,000,000 to 10/1,000,000 second (1 to 10 microseconds).

Computer Memory

Even small computers contain several types of memory. Internal computer memory is that which is designed into the computer. When questioned, it always provides the same, unchanging, predictable

information. In this memory bank are found special instructions to the computer, computer language, and number tables. This category is called **read only memory** (ROM). It is normally permanent. One special form is known as erasable program read only memory (EPROM). With special equipment its content can be modified or changed.

The other basic category of memory is used for temporary storage of instructions (programs) and of data to be processed or being worked on. This memory bank, called **random access memory** (RAM), must have a large capacity since a sizable amount of instructions and data must be stored. (The computer must first look for stored data and instructions before it can process them.)

Magnetic Core Memory. Because materials that contain iron, either in its pure form or as part of compounds or alloys, can be magnetized in two directions, binary numbers can be represented by using small beads of ferrite (finely divided iron) magnetized by passing a DC current through the wire on which they are strung. This approach was one of the earliest methods of storing data (numbers) in a digital computer. Since the cores remained magnetized even when the machine was turned off, the system was called core memory (Figure 13.4).

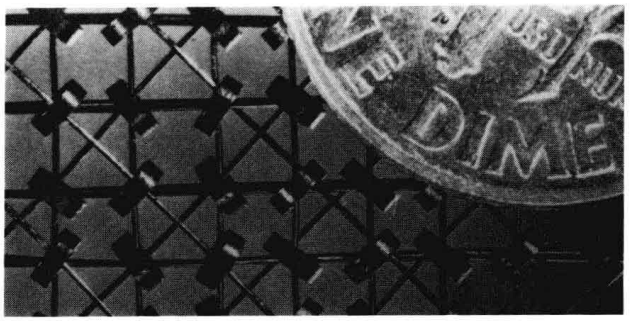


Figure 13.4 Magnetic Core Memory In the early 1970s ferrite iron cores strung on tiny conducting wires were widely used for storing and recalling computer data. Data was stored in the form of a changed polarity of the magnetic field surrounding each ferrite core. (Photo courtesy of General Electric)

IC Memory. Core memories were difficult to construct and were much bulkier than integrated circuits having the same number of elements. Although IC memories are ultracompact and provide quick access to desired data, they have one serious drawback: If the power fails or if the operator throws the wrong switch, the entire contents of the computer's memory may be forgotten. A memory that can be lost in this manner is said to be *volatile*. Loss of memory can be prevented through the use of standby batteries that are automatically switched into the circuit if the power fails.

Dynamic Memory. A recent improvement in IC memories, dynamic memory consumes only about one-fifth as much DC power as earlier types. Instead of constantly drawing power as transistors do when they are continuously turned on, the dynamic memory has its power applied in short bursts that last only long enough to maintain the IC in a charged condition. Typically, the charge must be replaced 500 times a second.

The term *dynamic memory* derives from the fact that while the memory transistor is on, the charge on its control element is changing. In contrast, earlier types of IC memory are static since the potential of one of their control elements remains unchanged.

The additional circuitry required to refresh dynamic memory is small compared to the advantages of being able to reduce by a factor of 4 in the number of transistors required for a certain amount of memory. Fewer transistors result in a significant cost reduction. Even more important, four times as much memory can occupy the same physical space.

How Memory Is Quantified. Since what a computer can accomplish is directly tied to the size of its memory, particularly its RAM, memory capacity is always mentioned in the sales literature. Computer memory is normally measured in **bytes**—single groups of **bits** processed together. The byte is much too small for convenient use, so in practice we use approximately 1000 bytes, or 1 **kilobyte** (1K). Actually, 1K represents somewhat more than 1000 bytes since in binary math 2^{10} is 1024. How-

ever, in normal use we consider 1K as 1000 bytes. As of early 1984 the minimum acceptable RAM capacity for a personal computer was about 64 kilobytes.

Memory Storage. *Magnetic patterning* is a convenient way to store instructions and data. The process is exactly like recording voice and music on an audiotape recorder. Magnetic tape recording of data is currently used with all sizes of computers, even personal computers that employ an audio cassette machine. However, magnetic tape recording involves having to wind through a reel of tape to locate the desired data.

To reduce the time required to access (find) data, *magnetic disc recording* was developed. Magnetic coating was applied to rigid discs of a nonmagnetic material, usually aluminum. Using a number of these discs (10 to 14 inches in diameter) together with multiple recording and playback heads reduced the time required to access specific data to a fraction of a second.

Because of the mechanical precision required, the rigid disc system was costly. Experimenters then tried using a disc of flexible plastic, which permitted a single record/playback head with a dual purpose to follow the disc with ease. Since a lesser degree of mechanical precision was required, manufacturing costs were reduced. The flexible plastic disc, now known as a **floppy disc,** has undergone extensive development. Originally 8 inches in diameter and with a memory capacity of 160K bytes, the floppy disc has been reduced in size to 3.5 inches and its storage capacity increased (through use of both sides of the disc) to 1 megabyte (1,000,000 bytes).

Through the use of IC technology, record/playback heads were miniaturized and physically lightened so that they could fly 30 to 40 millionths of an inch above the surface of a rigid disc. No physical contact occurred between head and disc. This technique reduced the precision required in hard discs. Their size was also reduced to 5.25 inches in diameter.

Under development for some time as an approach to magnetic recording of data has been *bubble memory,* which uses a branch of IC technology to replace the moving parts of magnetic disc drives with a thin film of deposited material. By forming miniature coils, chains of magnetic fields move in a circular path within a thin layer of magnetic material. When viewed under a microscope, these fields appear as bubbles that move. Bubble memory is part of a new, portable computer in which it takes the place of a conventional, low-cost disc drive of 128K bytes capacity. Although currently about twice as expensive as a comparable disc drive, the bubble memory's small size (about 2 inches square) makes it attractive to design engineers, who expect its present price to fall as production increases.

Business firms, in particular, may require that large amounts of data ready for quick retrieval—for example, personnel records, tax payments, and inventory control records—be stored for periods of one to five years. Floppy discs offer the advantage of a large amount of data storage surface plus quick access to the right information by the computer. However, if an enormous amount of data is to be stored, the best answer still seems to be a reel of magnetic recording tape. Although a reel of tape can accommodate a tremendous amount of data, the computer must search out the location on the tape where a particular piece of data is stored.

Input-Output Circuits

The computer operator must have means to communicate with the machine. Such facilities are known as **input-output circuits** (I/O). The I/O circuits serve two functions. First, they make it possible for instructions and data (in proper form) to be fed into the computer and for processed data (again in proper form) to be recovered from its output. The various equipment items connected to the input and output terminals—for example, electronic keyboards, floppy disc readers, tape readers, magnetic tape cassette players, and cathode-ray tubes—are known as **terminal equipment** (Figure 13.5).

The second function performed by the I/O equipment is that of permitting the signal (data in

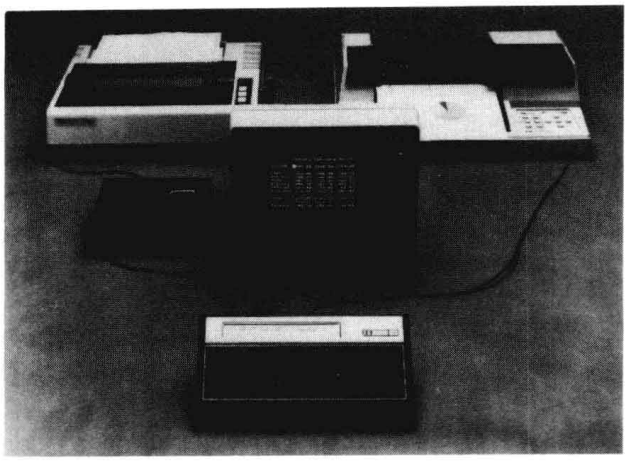


Figure 13.5 Terminal Equipment Various types of terminal equipment, or peripherals, are connected to a small computer (lower center) to provide additional facilities. Shown is a printer (upper left), a plotter (upper right), and a video monitor (center). The small black box is a tape drive that provides additional memory. (Photo courtesy of Hewlett-Packard)

proper form) to pass freely from the terminal equipment to the computer and from the computer to the terminal equipment. The I/O equipment must also isolate the computer electronics from the electronics of the terminal equipment. This function is referred to as **buffering.** The devices that perform these functions on the I/O circuits are known as **interfaces.**

Busses and Computer Language

Not actually a computer building block, but absolutely essential to its operation, are the electronic highways, or **busses,** over which instructions to the computer and data for it to process are constantly flowing. A bus is a conductor used as a path over which information flows from any of several sources to any of several destinations. Much of the information is in the form of 8 bit and 16 bit "words." This information takes the form of pulses of DC voltage of extremely short duration. These data pulses also flow back and forth among the three major operational blocks of the computer and to the output terminals.

Traditionally, computers have been designed around a permanent network of busses made up of many interconnecting conductors. This interconnection is referred to as **hard wiring,** which can accommodate many binary numbers (ones and zeros) in a single, horizontal line of data. Each individual digit of this line carries one bit of information. An entire line of 4, 8, 16, or more such bits provides one byte of data.

The byte can be used as the equivalent to one letter of the alphabet. By using two or more such letters together, we establish a mnemonic code as a basis for a computer language. (The word *mnemonic* is Greek in origin and means "memory aid.")

Throughout the brief history of computers, there has always been a need for a **computer language** that would permit nontechnical people to talk to a computer. The idea common to all the languages developed is to substitute a brief word that the operator can remember and that can convey information in digital form (both instructions and data) to the computer. Circuitry within the computer must first translate the English word to a string of ones and zeros, the only language to which the computer can respond.

Scientists were first to develop FORTRAN (FORmula TRANslator). Business people soon followed with COBOL (Common Business Oriented Language). The recent widespread interest in microcomputers has spurred the use of BASIC (Beginner's All-purpose Symbolic Instruction Code). BASIC is particularly helpful to the hobbyist since the use of such words as LET, LIST, GOTO, PRINT, RESTORE, RUN, and STOP results in actions by the computer that parallel the words' meanings in English. Through the use of BASIC the operator can address the computer's input in meaningful English words easily recalled.

Hewlett-Packard's HP-87 personal computer features a high-resolution, integrated 80 column text and graphics display, enhanced HP BASIC programming language, expanded software support, and memory that can be increased from 32K to 544K bytes. It approaches the capabilities of much larger computers. (Photo courtesy of Hewlett-Packard)

Home Computer Installations

So many personal computers are being offered at competitive prices that prospective buyers should first pick a reliable dealer to recommend a model that will best meet their needs. The dealer should also be able to service the computer if problems develop. Prices have fallen dramatically since the shake-out of mid-1983. However, any computer chosen should have at least 64K of RAM memory.

Essential Peripheral Equipment

The **keyboard** is technically a **peripheral** piece of equipment, though usually packaged as part of the computer. Next in importance is the video monitor. While a TV set can be used as a monitor, a great deal of definition will be lost. A monitor will provide far better resolution that enables you to read extended columns or lines of numerals with ease. The CRT screen should have a tube with a phosphor coating that glows in either the light gray or green part of the spectrum rather than in the blue or blue-white range. (Tubes in the blue range cause excessive eye fatigue.)

Optional Peripheral Equipment

Because so many computer programs (sets of instructions for the computer) are only available on disc, a relatively inexpensive **disc drive** is a valuable addition to a home computer installation. The disc drive can also be used to expand the RAM and data-storage capabilities of a personal computer. (In this application the program disc is removed from the disc drive and a blank magnetic disc substituted.) Retail prices of disc readers have been falling as production has increased, but they still cost as much as computers themselves. Discs designed for one computer will not work on the disc reader of a different computer.

A **dot-matrix printer** makes it possible to record data in printed form. The printer's name derives from the fact that the characters it produces on paper are made up of myriads of tiny, inked dots in a closely controlled pattern. The dot-matrix printer should not be used for word processing since the quality of the printing is not up to commercial production standards. A daisy wheel, or impact printer, produces higher-quality printing but is far more expensive.

Communicating with Other Computers

Operating a personal computer can be more rewarding and interesting if you can use it to communicate with other computers. The most promising approach is to use ordinary telephone lines as a communications link. However, the output of a computer is digital information, and a telephone can handle data only in analog form. This problem is resolved by an electronic device called a **modem** (MODulator/DEModulator) that sells for about $150. A mo-

dem uses the latest IC circuitry to convert the digital output of a computer to a series of audio tones that fall within the comparatively limited audio frequency range the telephone system can handle.

New Computer Offerings

The number of personal computers in the United States grew rapidly until early 1983, when a marked downturn occurred. Currently there are 4 to 5 million PCs in U.S. homes. Although stocks of PCs are moving slowly despite sharp price cuts, buyers have focused on the higher-priced lines—for example, the IBM PCjr and the Apple Macintosh—which offer a wide range of **software** covering both operations and subject material. (Software is a program of instructions that tells the computer what to do.) With suitable terminal equipment, these computers sell for well over $3000.

The conventional desk-top personal computer, with its full-size keyboard and CRT display, has recently acquired a vigorous competitor: the pocket computer. It is small enough to fit in a pocket, is battery powered, and costs under $100. Its keyboard has been miniaturized, but its keys are conventionally located. A separate keyboard row is used for numeric and function keys. Its display is made up of a panel of liquid crystals (LCDs) rather than the usual CRT. The display capability is limited by its size rather than by the capabilities of the computer itself. If more data are needed than the 24 character crystal display can provide, controls permit scrolling the data in the computer so lines of data above and below the original line can be seen. Both RAM and ROM memory capacity is less than for conventional desk-top machines but is adequate for many applications. Because the unit is battery powered, it should be turned off promptly when not in use and the battery recharged. A full line of matching peripherals, such as printers, will soon be available. A wide offering of software programs is available in BASIC.

An analog to the transistor, which employs electrons in its operation, is under development. The new device uses laser beams directed through crystals whose refractive index varies with the intensity of the light. One of the factors that limits the operating speed of today's computers is the time required for electrons to pass through the semiconductor material of the many transistors required. If laser beams could take the place of electron flow, it seems possible that a computation could be made in one-billionth of a second. Many years of work will be required to produce the first operational model of such a computer.

Hewlett-Packard's HP-75 hand-held computer performs most of the functions of a conventional desk-top machine. Weighing only 26 ounces, it easily slips into a pocket or briefcase. The display is liquid crystal and can be moved up or down to reveal additional data. (Photo courtesy of Hewlett-Packard)

MICROPROCESSORS

The microprocessor is closely related to the digital microcomputer. Both are designed around a central processing unit mounted on a single silicon chip. If the CPU is very sophisticated, the addition of a memory unit and I/O terminals results in a microcomputer. However, microprocessors are not designed with the same objectives as computers. Consequently, the CPU of a microprocessor is not very sophisticated, the amount of memory incorporated

is minimal, and the input-output circuits are designed for precise control over some process or operation.

In other chapters we discuss how microprocessors exert control over many operations—for example, the adjustment of f-stop and shutter speed to ensure proper film exposure in a camera, the temperature range and cooking rate in a microwave oven, and the temperature and drying time of a home laundry unit. Microprocessors also have applications in the automotive industry. By 1982 new cars coming off automotive production lines had at least two microprocessors—one to control the ignition system, the second to supervise operation of the fuel system, including fuel metering and injection. Microprocessor use will also be expanded on automotive and other production lines to maintain quality control through constant checking of components and subassemblies for errors in assembly and adjustment. Cars manufactured in 1983 use additional units for speed controls and in voice-synthesis modules. (See Chapter 16 for more automotive applications.)

Microprocessors of the mid-1980s are far more sophisticated than those produced four or five years ago. They usually have more memory, and some have both analog and digital outputs. Digital output allows the simultaneous use of the microprocessor as both a control device and a driver for an LED or LCD display. Newer microprocessors can also be used to supervise and coordinate the operations of other microprocessors.

ELECTRONIC CALCULATORS

The widespread use of electronic calculators stems from the discovery of the technology, a decade or so ago, of constructing by a photochemical process a complete CPU on a silicon chip less than 1/6 inch square. Mass production of millions of these units, the most costly item in previous calculators, caused prices to tumble dramatically. Better calculators can be purchased today for far less money than was the case a dozen years ago.

Hewlett-Packard's 1983 version of its popular HP-92 portable printing calculator is designed to analyze the return of various investments in bonds and real estate mortgages. It can also be used to handle other mathematical and statistical data. (Photo courtesy of Hewlett-Packard)

Construction of Electronic Calculators

It is convenient to think of the functioning of an electronic calculator in terms of the same building blocks we used in studying computers and microprocessors. The calculator is simpler in design than the microprocessor and far less complex than the microcomputer. Its CPU can therefore be less complex, which greatly lowers its unit cost.

One important function of the CPU is to perform the basic arithmetical operations (addition, subtraction, multiplication, and division) by variations on the basic process of addition, just as the computer does. Before it can start the addition process, the numbers punched in from the keyboard must first be converted to 4 bit binary coded decimals because as in the computer, the data to be worked with can only be in the form of a series of ones and zeros. A circuit incorporated in the CPU chip automatically performs

this conversion when the number key is depressed on the keyboard. A somewhat similar circuit converts the binary answer back to analog form for the display.

The calculator's keyboard corresponds to the input terminal of the computer. Its display corresponds to terminal equipment connected to the computer's output terminal. The keyboard is very simple—merely a group of off-on switches, most assigned to numerals and some used to select the desired function.

Trends in Visual Displays

Early calculators used **light-emitting diodes** (LEDs) to provide an 8 digit display. By dividing each LED unit into seven segments (Figure 13.6), all the digits from 0 through 9 can be produced by proper switching. Since the reddish glow given off represents power generated, there is a comparatively heavy drain on the 9 volt battery that powers calculators using this type of display.

In recent years calculator manufacturers have focused on **liquid crystal displays** (LCDs), which offer decided advantages. They employ certain carbon-based organic compounds that simultaneously behave like crystals and liquids. When subject to an electric field, their molecules tend to line up in one pattern. When the field is changed, they form another pattern. This change in patterning also causes them to reflect or absorb light selectively. They do not generate light; they merely reflect and absorb light that falls on them. Since light is not generated, power consumption is reduced to perhaps one-tenth of that required for an LED display. Therefore, silver oxide batteries of the type used in cameras and watches can be substituted for the much bulkier carbon-zinc types. It is now possible to produce an ultrathin calculator (less than 1/4 inch thick) and at the same time to greatly extend battery life. In many cases new batteries have operated well over a year before replacement was necessary. In the mid-1980s, research continues on new battery components that will provide still longer performance. Solar-powered units have also been developed.

Figure 13.6 LED Display The 7-segmented LED or LCD can form all letters and numerals.

Trends in Calculators

Besides the trend toward compactness is a move in the direction of specialized calculators—ultrasimple calculators for children in the elementary grades, scientific calculators for college majors in the physical sciences, and calculators for realtors to compute declining interest payments on a mortgage as the principal is paid off. Calculators for business executives and accountants record all totals and subtotals on a roll of chemically treated paper. A stylus, moved by a basic microprocessor, traces the numerals on the moving strip of paper, which is treated to change color when a DC current passes through it.

By 1983 the better calculators (and some very low cost ones) included function keys that provided access to a completely addressable memory bank. The memory bank permits storage of subtotals before the rest of the data has been processed. It can also be used for storing a corrective factor that is to be used many times in a series of computations.

CLOCKS AND WATCHES

People have always been preoccupied with the passage of time—not only the passage of hours but also the waxing and waning of the seasons. Ingenious means of measuring time have been developed. The sun dial in a primitive form has been in use for tens

of thousands of years, but it is worthless on a cloudy day or at night. The water clock, which measured the hours in terms of how long it took a certain amount of water to leak through a tiny opening, apparently originated in China but spread to Egypt, Greece, and the Roman Empire. Crude mechanical clocks originated independently in Germany, France, and Italy as early as the 13th century, but the first pendulum clocks were not seen until about 1700. Some of these clocks included minute and second hands.

In a mechanical watch or clock the time-keeping function is primarily carried out by the oscillation of a balance wheel attached to a balance spring (hair spring). A modern watch of this kind is capable of an accuracy of ±30 seconds in a 24 hour day, a degree of precision at one time considered remarkable.

In the early 1930s the electric clock appeared. It was basically a series of gears driven by a synchronous electric motor slaved to the 60 cycle current. The gear ratio was such that the hour, minute, and second hands swept the clock face and its numerals at the proper speed. Timekeeping was as accurate as the accuracy of the power company's generating equipment.

In this section we discuss purely electronic clocks and watches—that is, clocks and watches having no moving parts. Many clocks and watches of a few years ago were advertised as electronic when in fact they were partly mechanical and partly electronic. Although they did incorporate electronic circuitry to improve their performance, they required periodic cleaning and oiling like purely mechanical types. Also, since they expended energy in producing motion, battery life was short unless the unit was powered from 110 volt power lines.

Concept of Timekeeping

The higher the frequency of the **time base,** or reference frequency, the greater the overall accuracy of the total system. In a conventional mechanical watch, the time base is 1/5 second since the balance wheel oscillates back and forth five times per second. In hybrid models, the time base was 360 hertz, the natural frequency of the tiny tuning fork used, and the accuracy was further enhanced since the base was increased by a factor of some 1800 times. For a while it seemed the industry was to standardize on 32.768 kilohertz as a crystal-controlled timing base, but today there are many deviations from this standard. However, with such a high timing base in use, a theoretical accuracy of ±30 seconds a year should be entirely within reach.

Construction of Electronic Clocks

The best way to learn how an electronic clock is assembled and functions might be to purchase a kit of parts and actually construct one. An Optoelectronics model 850-4 includes such options as 50 or 60 hertz time base, 12 or 24 hour time, and the choice of the line frequency or a crystal oscillator of high accuracy as the time base. Figure 13.7 shows an overall view of the 3 1/2 by 4 inch circuit board and components. Not shown in the photo are the power transformer, which is molded into the power plug (110 volts AC to 12 volts AC); the three toggle switches used to adjust the clock time initially; and the crystal oscillator circuit, which is first constructed on a separate piece of circuit board.

Reading from the upper surface of the circuit board, the first row of components is made up of six LEDs (light-emitting diodes) capable of displaying readings in hours, minutes, and seconds. The next row contains the current-limiting resistors for the emitter circuits of the six driver NPN transistors in the row immediately below. They amplify the output of the timing pulse from the clock chip to a level that will produce full illumination of the LEDs. An extra transistor is provided for adding a colon between the hour and minute digits or between the minute and second digits.

The most important single component lies in the next row: the National Semiconductor MM-5314 clock chip IC. It has a total of 24 pins (12 on each side), not all of which are used in this particular circuit.

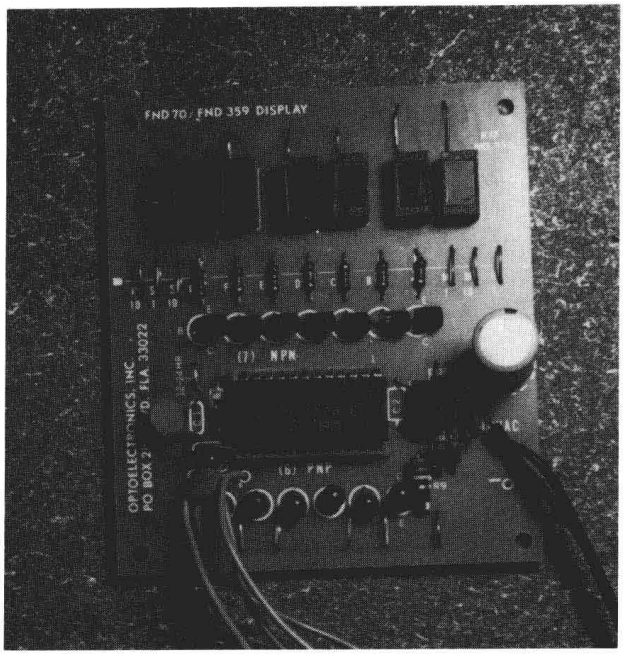
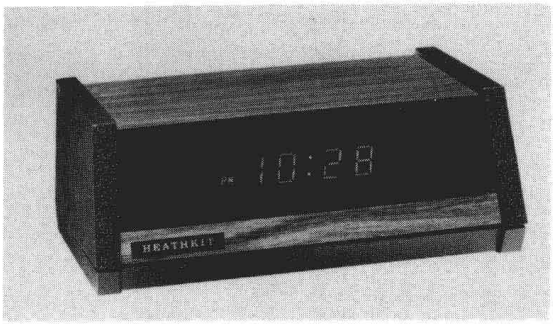


Figure 13.7 Digital Clock Unit The circuit board of an Optoelectronics clock unit includes (from the top) 6 LEDs, 7 NPN driver transistors, the IC unit providing the timing pulses, and the 7 NPN isolating transistors.



Heathkit's model GC-1107 digital clock, intended for operation on 110 volt AC power, has a brilliant LED digital display. It also provides an alarm function. (Photo courtesy of Heathkit)

The clock chip IC is the equivalent of hundreds of separate transistors, capacitors, resistors, and their associated wiring. just to its right are the four diodes used to change the 12 volts AC from the transformer into 12 volts DC required for proper operation of the transistors and the integrated circuit. Just to the right of the diodes is a large electrolytic capacitor that smoothes out the slightly pulsating DC from the diodes into pure DC. The final row consists of six PNP transistors that are used to interconnect the pins on the clock chip and the driver transistors. They provide isolation between the two circuits and additional amplification.

Crystal Control Circuit

The basic clock just described can be vastly improved in accuracy and its usefulness can be widely extended by the simple addition of a crystal-controlled time standard intended for easy installation in the basic clock kit. The two most important components of the time standard are the quartz crystal (manufactured for color-burst applications in color TV sets) and the special MM-5369 integrated circuit from National Semiconductor. With these components the clock we have described has an accuracy measured in terms of seconds per week or month rather than minutes per day.

After disconnection of the power transformer, the clock is operational from a 12 volt DC source such as an automobile battery, so its usefulness is greatly extended. Most high-quality digital clocks employ crystal control and frequency divider circuits.

LED and LCD Displays

Most digital clocks sold today are intended for household or automotive use, where the comparatively high power consumption of LEDs poses no problem. Not all digital clocks consume a great deal of power. RCA uses a 7-segmented miniature incandescent lamp to form each digit, while other manufacturers are turning to liquid crystal displays (LCDs) whose lower power drain permits powering a clock

for perhaps a year with a single heavy-duty cell. Since LEDs actually generate light, a large display can be read across a brightly lit room. Because LCDs can only selectively reflect and absorb existing light, they are far less effective in bright light.

Electronic Watches

Practically everything we have said about electronic clocks is applicable to electronic watches. Their development had to await that of quartz crystals, integrated circuits, and long-life batteries of extremely small dimensions because bulky, awkward watches could not be sold. A decade ago a mechanical wristwatch that did not gain or lose more than 30 seconds a day was considered acceptable. Today's high-quality digital watch can be factory adjusted to maintain its accuracy within 30 seconds a year.

Because of the high current drain of LED displays and the additional electronics they require, LCD displays are almost always used in today's digital watches. Improved design of the module through the use of field-effect liquid crystals and optical filters has resulted in a display that is much easier to read, particularly at an angle. The low current drain makes it feasible to operate the display continuously rather than at the push of a button. Many watches include data beyond the time of day—for example, the current month and day. A digital watch can share its case with a miniaturized electronic calculator (Figure 13.8). The two units are fully interconnected and include a 200 year calendar, a precision stopwatch, an alarm clock, and guaranteed watch accuracy of ±30 seconds per year. Some watches have built-in calculators or photovoltaic cells to maintain the charge of an internal storage cell.

Figure 13.8 Calculator-Watch This 8 digit, 4 function calculator, combined with a quartz digital watch, features a calendar, dual time, and a 24 hour electronic alarm. It can also be used as a stopwatch accurate to 1/100 second. (Photo courtesy of Radio Shack, a Division of Tandy Corporation)

SUMMARY

People have always sought ways to make simple calculations easier. The first calculating machine, the abacus, was based on base five math. Twenty centuries later the next computing machine was developed by Charles Babbage. His approach was purely mechanical, but his machine incorporated many of the basic principles of today's computers. In 1944 IBM completed the world's first electromechanical computer. The first computer to use vacuum tubes was soon in operation at the University of Pennsylvania. It was much more rapid than the first IBM machine and underlined the importance of active devices in computer design. With the development

of the transistor in the early 1950s, computers grew dramatically smaller and their speed of operation increased. Intel's development of an integrated circuit containing a complete central processing unit on one silicon chip made possible the development of the microcomputer at remarkably low cost by the late 1970s.

All modern computers are designed around base two math, which permits all numerical data to be expressed in just two quantities, ones and zeros. In performing the basic arithmetical functions, the computer in effect always performs addition. Addition is straightforward, subtraction is addition with the sign of one quantity changed, multiplication is essentially sequential addition, and division is essentially sequential subtraction.

To store instructions and data, the computer can use magnetic cores, magnetic tape, floppy discs (with magnetic coatings), or ICs. The computer must also be equipped with input/output circuits so that data and instructions can be fed into the unit and processed data retrieved.

Today's calculators, with their expanded memory capabilities, are becoming more like computers. The keyboard is the input circuit; the display is the output circuit. As with the computer, the calculator functions are all based on addition, but mathematical shortcuts may be used. Newer calculators are much smaller than those of a few years ago because liquid crystal displays (LCDs) with their very low current drain make practical the use of dime-sized watch or camera batteries that last at least a year.

It is easy to confuse electric clocks and electronic clocks. The familiar electric clock is powered by a synchronous electric motor slaved to the 60 hertz frequency of the power line. Its accuracy is as good as the power company's ability to maintain the frequency of its AC generators at precisely 60 hertz. An electronic clock may operate entirely independently of the AC line frequency, though it may be powered either from a DC power supply connected to the line or from a battery. Its time base comes from an oscillating quartz crystal connected to a special IC that takes the crystal's comparatively high frequency (several megahertz at least) and divides it down to exactly 60.000000 hertz. Other circuits actuate the numeric digit display to indicate the correct time. Electronic watches, although they use much smaller components, work on the same principle and are built to higher standards of timekeeping. A well-designed watch will not gain or lose more than 30 to 50 seconds in an entire year.

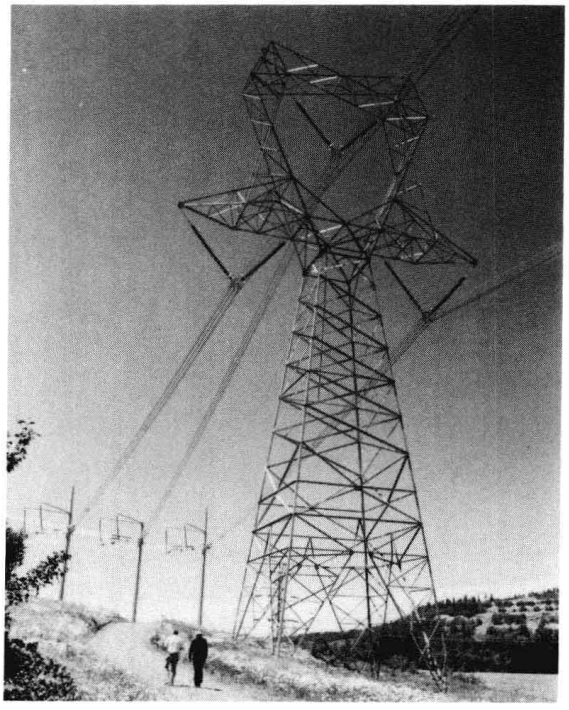
KEY WORDS AND CONCEPTS

abacus
analog computer
arithmetic logic unit (ALU)
base two math
binary numbers
bit
buffering
bus
byte
central processing unit (CPU)
computer language
digital computer
disc drive
dot-matrix printer
floppy disc
gate
hard wiring
input/output circuit
interfaces
keyboard
kilobyte
LCD display
LED display
microcomputer
minicomputer
modem
module
peripherals
program
random access memory (RAM)
read only memory (ROM)
software
terminal equipment
time base

SELF-TEST QUESTIONS

1. How does an analog computer differ from the more common digital computer?
2. Why did early digital computers require elaborate air conditioning systems?
3. Engineers of the Intel Corporation are generally credited with one of the greatest advances in computer design in the 1970s. What was it?
4. What are some of the many functions readily performed by a modern computer?
5. Explain how all of a computer's mathematical functions are linked to addition.
6. The conventional math is base ten. Why does the use of base two math simplify the design of the computer?
7. What are the three basic building blocks of any computer?
8. What three types of magnetic memory storage are in common use?
9. What new electronic development promises to replace magnetic memory for temporary data storage?
10. Why is some form of computer language always needed?
11. How does the basic design of a microprocessor differ from that of a digital microcomputer?
12. In a pocket calculator, what corresponds to the input/output circuits of a computer?
13. What type of transducer is used as the time base of the most accurate electronic clocks and watches?
14. What recent development in electronic watches has extended battery life and made possible a continuous digital display?

14 Generation and Distribution of Electrical Power



This experimental three-phase distribution line operates at the unusually high voltage of 1200 kilovolts (1.2 million volts). Conductors are connected in parallel to reduce the transmission line resistance. (Photo courtesy of Bonneville Power Administration, USDOE)

Historical Perspective
- First Power Station
- Introduction of AC Power Generation

Basic Concepts of Power Generation
- Power Generation
- Transmission Lines
- Typical Power Distribution System
- Steam Plants
- Problems of AC Power Distribution

Distribution of DC Power
- Principles of DC Transmission
- Converting DC Back to AC
- Future Trends

Nuclear Power Generation
- Atomic Fusion
- Atomic Fission
- Nuclear Power Plants

Alternate Sources of Electrical Energy
- Solar Energy

Wind Power
Metglas and AC Power Distribution

Summary

Key Words and Concepts

Self-Test Questions

In this chapter we trace the development of the equipment and the technology necessary for the generation and distribution of electrical power. At first electrical power was used primarily for electric lighting. As heavier equipment became available, it was used for the production of mechanical energy. Today electrical power finds additional application in energizing an enormous variety of electronic devices.

At the turn of the century, the most convenient source of energy to spin the shafts of the large generators in use was steam from coal-fired boilers. As the technology of transmitting power over a distance by a transmission line was worked out, power produced by falling water at hydroelectric sites proved more economical. To supplement the output of hydro power at peak periods of consumption, steam plants burning crude oil were constructed in the major cities. Some urban areas remote from easy access to hydroelectric power (e.g., Los Angeles) placed nearly complete reliance on oil-fired installations.

Nuclear power generating plants closely resemble the steam-turbine plants fired by oil or coal. The basic difference is that the heat source in a nuclear plant is a nuclear reactor that produces heat by a process called fission, which is explained in a later subsection.

In the mid-1980s the so-called fossil fuels are in improved supply, but interest is still high in alternative electrical power sources, the most promising of which seem to be the radiant energy of the sun and the energy of the earth's prevailing winds. In a later subsection these power sources are explored in detail.

Finally, we stress throughout the chapter that generation and distribution of electrical power is a technology that every year makes additional use of electronics at an increasingly sophisticated level. The past decade has seen a large-scale trend toward replacing AC power transmission lines with DC lines—a move that was delayed at first by the lack of suitable controlling electronic equipment to withstand the high DC voltages encountered.

HISTORICAL PERSPECTIVE

The paradox of public utility engineers returning to DC for high-voltage transmission after abandoning it nearly a century ago in favor of AC can be better understood after a review of the electrical power industry's history. We cannot describe the social and economic revolution that low-cost, readily available electrical power made possible without mentioning the contributions of Thomas Alva Edison, the "wizard of Menlo Park." Although Edison, contrary to popular opinion, did not invent the electrical generator—Michael Faraday had established its working principles as far back as 1831—he did improve its effectiveness and reliability. Note: What we today refer to as **generators** and **alternators** were universally referred to as **dynamos** in Edison's time.

An eminently practical man, Edison made two important decisions that would affect the course of electrical power generation and transmission for over twenty years. The first decision was that DC power was more efficient than AC power since in the power equation $P = E \times I$, with both voltage and amperage constant at all times, the amount of power transmitted would be maximum. His second decision, also based on practicality, was that 100 volts was a good value for the generator voltage. Higher voltages introduced problems of arcing and insulation breakdown, while lower voltages required heavier line currents to transmit the same amount of power over the system. Later, as generator construction improved, Edison experimented with higher transmission voltages.

First Power Station

Edison was anxious to construct a model electrical power generating station and distribution system. His first working model, the Pearl Street station, served 85 New York City customers with a single generator. Two years later, in 1884, he had a backlog of customer applications and was serving over 500 customers, including both residences and small businesses.

Since DC voltages gradually dropped because of the resistance losses in the transmission lines and could not be stepped up again, only an area a mile or so from a generating plant could be effectively served. As additional customers clamored for service at more remote locations, it became necessary to construct new generating plants and distribution systems to serve them. Soon Edison and his financial backers had a heavy investment in DC power plants in New York City and other urban areas. (Even today public utilities bearing his name are operating in New York and Chicago.)

Introduction of AC Power Generation

By the late 1880s George Westinghouse was challenging Edison's dominance in the electrical power distribution business. Having made a fortune from the sale of his patents on an air-brake system for freight and passenger trains, he was in a financial position to hire the best engineering talent to develop an AC system that he considered vastly superior to Edison's.

A key item in the AC power distribution system was the power transformer. By varying the ratio of primary to secondary turns, voltages applied to the primary winding could be stepped up or stepped down to the precise level required. Although Faraday had discovered the principle of the transformer in the 1830s, it remained for such gifted engineers as Tesla and Steinmetz to develop it to the point where it could be used in elaborate power distribution systems based on AC current. At the 1893 World's Fair in Chicago, the entire lighting system used AC designed and installed by Westinghouse engineers.

With characteristic stubbornness, Edison continued to cling to DC generation and distribution of power despite AC's ability to transmit large amounts of power at high voltages and low currents over great distances and then to step this high-voltage, low-current power down to the level required for most effective application. Under pressure from his financial backers, Edison finally conceded that power systems of the future would indeed be of the AC type. He even made a few contributions of his own. For example, the three-wire AC transmission system is an Edison development. He was also the first person to install and operate successfully an underground AC power distribution system in an urban area.

By the turn of the century, the trend toward AC as the preferred mode of electrical power generation and transmission was well under way. Since then the quality, effectiveness, and reliability of generating and transmission equipment have improved, but the basic technology has not changed materially. Only within the last decade or so have engineers considered returning to the high-voltage DC transmission of power for increased system effectiveness.

BASIC CONCEPTS OF POWER GENERATION

Well over 90% of the world's electrical power is produced electromagnetically, while only a comparatively small amount is generated by other means, mostly by chemical reactions taking place in cells and batteries. Only a minuscule amount is supplied by devices such as photoelectric cells and thermocouples. We discussed the generation of electrical power in Chapter 4. In this section we review briefly some of the basic concepts of AC and DC electrical power generation.

Power Generation

When you move a conductor through a fixed magnetic field or move a magnetic field past a fixed

conductor, a voltage is induced in the conductor that can cause a current to flow in an external circuit. This principle underlies all power generation.

Many things can be done to make the basic device produce higher voltages and higher currents. For example, you can form the conductor into coils that are wound over soft iron cores for maximum magnetic efficiency, you can increase the number of turns of wire in each coil, you can increase the intensity of the magnetic field, and you can increase the speed at which the coil moves through the magnetic field.

All of these changes are bound to reflect engineering compromises and trade-offs. You have to consider physical size of the component parts and their cost. You must also consider the rotational speed of the armature. Otherwise, centrifugal force will tear the entire generator apart.

Prime Power Sources

Generally the prime mover used to turn the shaft of the generators or alternators in major power installations is either **hydroelectric power** (falling water) or high-pressure steam. Nuclear plants fall into the second category, and the nuclear reaction is not used to produce electricity directly, but indirectly by providing heat to turn water into steam.

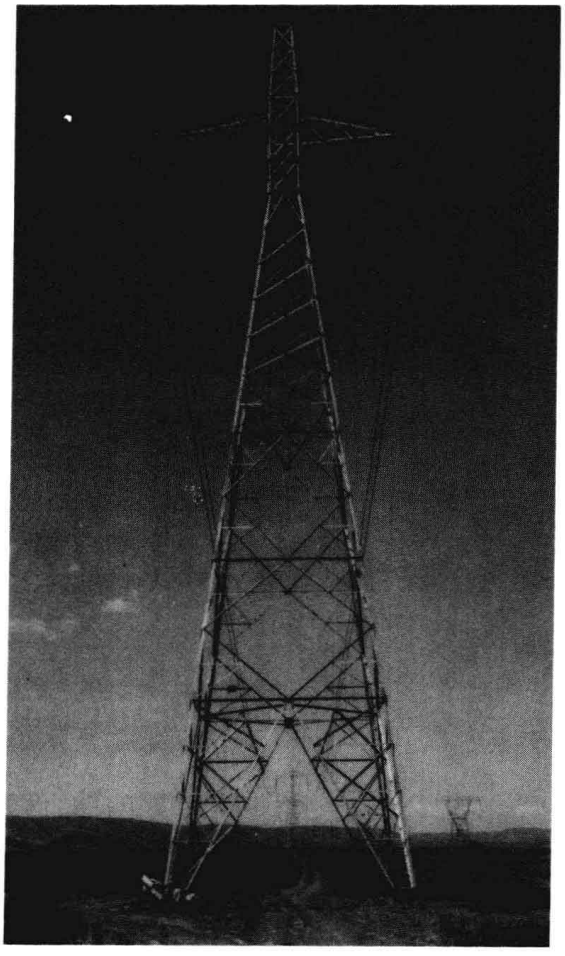
Most of the nation's steam-powered generating plants still rely on the heat generated by the burning of the prime **fossil fuels,** coal and petroleum. With the escalating costs and diminishing supply of crude oil, many of these plants, with government encouragement, are converting their furnaces to burn coal.

Varying Current Cycles

While 60 cycle (60 hertz) current is almost universally used in the United States, Canada, and Mexico, it is seldom encountered elsewhere. In Western Europe 50 hertz current is in universal use. To add variety, the power outlets in many hotels supply 220 volts rather than the conventional 110 volts at 50 cycles. When a plug-in adapter (actually a miniature step-down transformer) is used, 60 hertz devices like electric razors or magnetic tape recorders can be used, but their operation may be sluggish and ineffective.

Transmission Lines

One of the most important engineering achievements of the 20th century has been the development of the **transmission line,** which transports enormous



On a typical high-voltage DC transmission line, one pair of conductors is positive, the other negative. Note the comparatively small right-of-way required for tower and line. (Photo courtesy of Bonneville Power Administration, USDOE)

amounts of electrical power for great distances with the very minimum in circuit losses caused by the resistance in the line. As shown in Figure 14.1, the principles involved are very simple. In Chapter 5 we saw that when a power transformer is used in the step-up mode (more turns on the secondary winding than on the primary), the secondary voltage is stepped up, and the secondary current is stepped down by the same proportion. The amount of power has not changed—only its form is different. It is now high-voltage, low-current power instead of the low-voltage, high-current power supplied by the generator.

Putting Transmission Lines to Work

Almost everyone has had a close-up look at a modern transmission line with its large-diameter cables supported by elaborate insulators from elaborately reinforced steel towers. These cables are actually aluminum tubes that carry the current. They are strengthened by an internal steel cable to provide the necessary mechanical strength. Transmission line voltages are very high—measured in hundreds of thousands of volts—since the higher the voltage, the lower the line current for a given amount of power

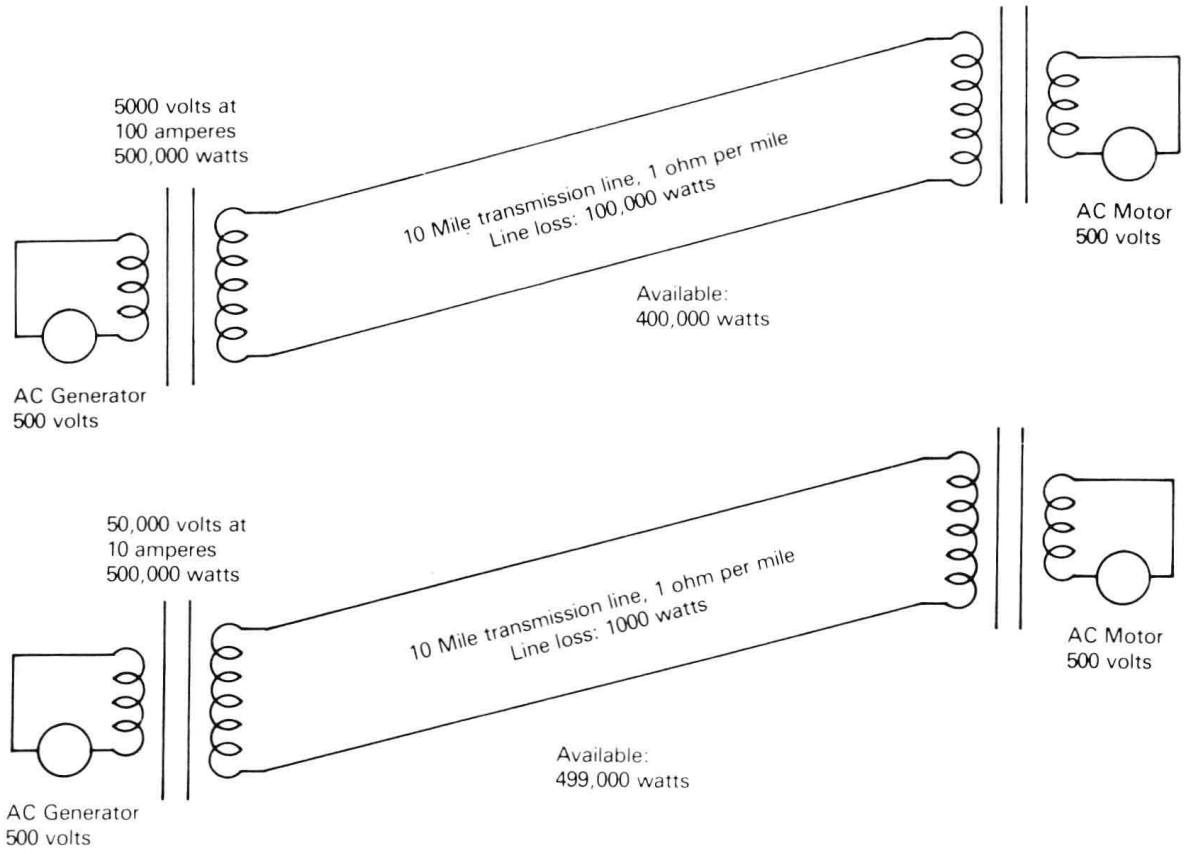


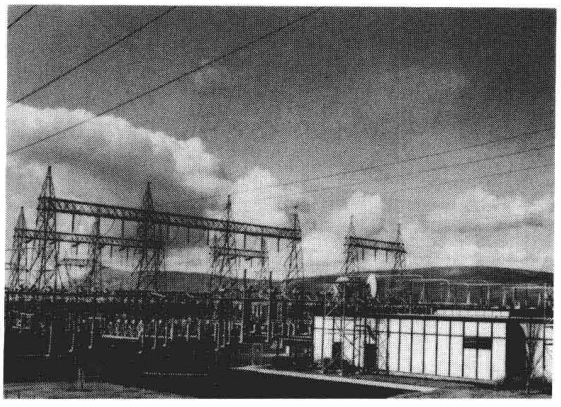
Figure 14.1 Schematic of Two AC Transmission Lines These lines transfer power from generating site to point of consumption. When the line voltage is extremely high, the line current can be quite low for a given amount of power transmitted in kilowatts or megawatts because the product of $E \times I$ is still high in the formula $P = E \times I$.

transmitted. In this manner the very minimum of power is converted into heat by the unavoidable resistance of the transmission cables.

Two Forms of Electronic Communication

Modern transmission lines also carry two important forms of electronic communications: **telemetering** (which means to measure at a distance) voice communications. By suitable technology, both transmitters and receivers can be coupled to the transmission line and used to transmit technical data as well as routine voice communications. The radio frequency energy generated by the transmitters used is superimposed on the 60 hertz line current and is not broadcast from a radiating antenna. Consequently, it can only be heard a very short distance from the transmission line, and even then it will be relatively weak.

Once the transmission line has delivered the energy from a hydroelectric site to an urban area perhaps 30 to 50 miles away, another very large power transformer must be used to step down the high-voltage, low-current power to a value of lower voltage and higher current suitable for distribution over the city's power grid. This transformer supplies the precise voltage and current values required by industries, businesses, and residences. The transformer and its auxiliary equipment is located in a fenced substation.



At the John Day substation, east of The Dalles, Oregon, incoming electrical power from a dam on the nearby Columbia River is transformed into higher voltages and lower currents for most efficient transmission. Protection against overload and switching are also provided. (Photo courtesy of Bonneville Power Administration, USDOE)

Typical Power Distribution System

An example of a typical power distribution system can be found in one West Coast city of about 100,000 population. Three hydroelectric sites are located 25 to 50 miles from town in the Cascade Mountains. They use the hydro energy of a sizable river of remarkably constant flow. Typically, the water-driven turbines spin at a controlled speed of 240 rpm, producing a generator output of 11.5 kilovolts (11,500 volts). A very large transformer is then used to take the output of a number of these generators and step it up from 11.5 kilovolts to 66 kilovolts (66,000 volts) to feed the three-phase transmission line. On the outskirts of the city, many miles away, this 66 kilovolts is stepped down by several transformers to a more suitable distribution voltage. Industrial areas are provided with about 12,470 volts for area distribution, and this voltage, in turn, is stepped down to 880, 440, 220, and 120 volts for special needs such as the operation of large motors or for electrical lighting.

In the business district, where most of the demand is for electrical lighting, transformers installed in sealed vaults below street level supply potentials of 120 and 208 volts (the latter primarily for elevators and air conditioning equipment). Another distribution circuit, in the form of a closed loop, rings the city at a potential of 12,470 or 7200 volts and feeds pole-mounted transformers that step this voltage down to 110 and 220 volts for residences and small businesses. The average home uses 110 volt current for all lighting fixtures and wall outlets; 220 volts for

electric ranges, water heaters, washer-dryer combinations, and perhaps air conditioning equipment.

This hydroelectric installation and its distribution system have been in operation without substantial change for at least three decades. A more recently constructed plant would differ primarily in that the voltage of the transmission line would probably be at least several hundred thousand volts, permitting a considerable increase in the amount of power transmitted with the same resistance losses. Additional power needs in the city mentioned are currently being met by the use of a standby steam plant and by purchases from the lines of the Bonneville Power Administration.

Steam Plants

The use of high-pressure, high-temperature steam to spin the rotors of AC generators is widespread since so many areas of high power consumption are far distant from suitable hydroelectric sites. Until the recent rapid rise of prices for fossil fuels such as crude oil and coal and for natural and manufactured gas, steam-produced electric power was quite competitive in price with power produced by falling water. Steam-produced power is now much more expensive but must continue to be used since most good hydroelectric sites have already been developed and are already providing their maximum electrical power generation. The newest steam generating plants are located so as to have ready access to a low-cost coal source—for example, in areas where new strip mining is being developed. (A special type of steam generating plant uses nuclear fuel as its heat source and is discussed later in the chapter.)

Although steam generating plants burning oil or coal as a heat source are becoming increasingly expensive to operate as the cost of fossil fuels continues to climb, they have one great advantage over hydroelectric systems: They can be quickly placed in service and as quickly shut down. Several periods of peak power consumption occur in any 24 hour day. One is from about 6:30 to 8:00 A.M., when electric lighting and electric-powered furnaces are turned on, electrically heated water is used for showering, and electric stoves are used as families prepare breakfast. A second peak period occurs between 5:00 and 7:00 P.M. and is further extended during the fall, winter, and spring months by increased use of lighting equipment as darkness falls. At both peak demand periods the local power utility, even if it owns its own hydroelectric plant, will probably fire up its standby steam plant and then shut it down at the end of the peak period. A modern oil-fired plant can be placed in operation in half an hour or less. A coal-fired plant takes much longer.

Problems of AC Power Distribution

Since the end of World War II, U.S. consumption of electrical power has increased yearly. The increase reflects the growth in population and in the number of electrical appliances in service. Many new appliances use far more electrical power than their counterparts of thirty years ago. A frost-free home refrigerator, for example, consumes at least 750 watts—more than twice the power drain of a standard unit of three decades ago. A coffeemaker consumes at least 1000 watts, far more than the electric percolator it replaced.

Within the last decade a host of new household appliances, all consuming sizable amounts of electricity, have appeared—for example, waterbed heaters and super-powered vacuum cleaners. Starting about 1960 and continuing to the present has been a period of rapid development and widespread sales of washer-dryer combinations, electric dishwashers, electric furnaces, and air conditioning equipment, all of which consume extravagant amounts of electrical power. The burgeoning sale of microwave ovens, with their average power consumption of 1500 watts, has added a further burden on electrical utilities.

To maintain system efficiency in the face of ever-increasing power loads, engineers raised the transmission line voltage (AC) to make possible the transfer of the maximum amount of power at the smallest possible loss caused by the resistance of the conductors. However, at a transmission line voltage

TYPICAL RESIDENTIAL ELECTRICITY USE AND COST, PACIFIC NORTHWEST (OCTOBER 1982)

Appliance	Typical Capacity in Watts	Cost per Hour at 3.4¢ per kWh	Typical Hours per Month Use	Monthly Cost for Energy Used	Gallons Water per Month for Generating Power
Coffeemaker	1000	$.034	25	$ 0.85	158,250
Dishwasher	1500	$.051	40	$ 2.04	379,800
Drill	360	$.01224	varies	varies	2,279/hour
Dryer, clothes	5000	$.17	30	$ 5.10	949,500
Electric blanket	150	$.0051	240	$ 1.224	227,880
Freezer	750	$.0255	200	$ 5.10	949,600
Furnace fan	400	$.0136	300	$ 4.08	759,600
Hair dryer	450	$.0153	4	$ 0.0612	11,934
Hand iron	1000	$.034	15	$ 0.51	94,950
Heater, portable	1250	$.0425	30	$ 1.275	237,375
Heating pad	55	$.00187	varies	varies	320/hour
Lights	100	$.0034	varies	varies	633/hour
Mixer	120	$.00408	5	$ 0.0204	3,798
Oven, regular	2000	$.068	varies	varies	12,660/hour
Oven, microwave	1500	$.051	varies	varies	9,495/hour
Radio-phonograph	175	$.00595	40	$ 0.238	44,310
Range, small burner	1250	$.0425	varies	varies	7,913/hour
Range, large burner	2100	$.0714	varies	varies	13,292/hour
Refrigerator, regular	300	$.0102	300	$ 3.06	569,700
Refrigerator, frost free	750	$.0255	300	$ 7.65	1,424,250
TV (color), solid-state	200	$.0068	150	$ 1.02	189,900
TV (color), tube	450	$.0153	150	$ 2.295	427,275
Toaster	1000	$.034	5	$ 0.17	31,650
Vacuum cleaner	420	$.01428	20	$ 0.2856	53,172
Waffle/sandwich maker	1100	$.0374	8	$ 0.2992	55,704
Washer, automatic	400	$.0136	25	$ 0.34	63,300
Waterbed heater	400	$.0136	250	$ 3.40	633,000
Water heater, 52 gal.	4500	$.153	100	$15.30	2,848,500

Note: Gallons of water per month for generating power are based on 6330 gallons of water to generate 1 kilowatt-hour of electricity. East Coast, Midwestern, and Southern power costs are much higher than shown here.

Source: Data from Bonneville Power Administration (USDOE).

of about 800,000 volts AC (800 kV), further efforts to increase the amount of power transmitted by more voltage become largely counterproductive. (To understand the technical reasons, review the sections in Chapter 4 dealing with inductance, inductive reactance, capacitance, and capacitive reactance.)

When an electrical current starts to flow through a conductor, it creates an electromagnetic field around the conductor that expands as long as the line current is increasing. This expanding field cuts the conductor and induces a voltage in it of such a polarity as to oppose the line voltage. Engineers have long referred to this phenomenon as counter electromotive force. Its net effect is to reduce the line current to less than its predicted value, which places limits on the amount of power that can be transmitted.

With an AC system the current during the course of one cycle rises to a maximum in one direction, decreases, and falls to zero. Then, reversing polarity, it rises to a maximum in the opposite direction and again decreases and falls to zero.

This opposition to changing values of transmission line current caused by the electromagnetic field surrounding the conductor is known as inductive reactance. As the transmission line voltage and current are increased in an effort to transmit more power, inductive reactance becomes increasingly important as a factor limiting the line's overall efficiency.

There are some practical restraints on the voltage that can be used effectively on an AC transmission line. Increasing the voltage requires increased spacing between the conductors, which in turn requires a larger supporting structure and a wider right-of-way for the transmission line. Also, insulators that will withstand tremendous voltages become very difficult to design and unusually expensive. Finally, the extremely high-voltage capacitors that engineers frequently use to neutralize in part the inductive reactance of the line are not only subject to breakdown (which can cause a system failure) but are also prohibitively expensive. Therefore, electrical engineers all over the world who were seeking a better approach to electrical power transmission returned to Edison's original concept of using DC rather than AC power transmission.

DISTRIBUTION OF DC POWER

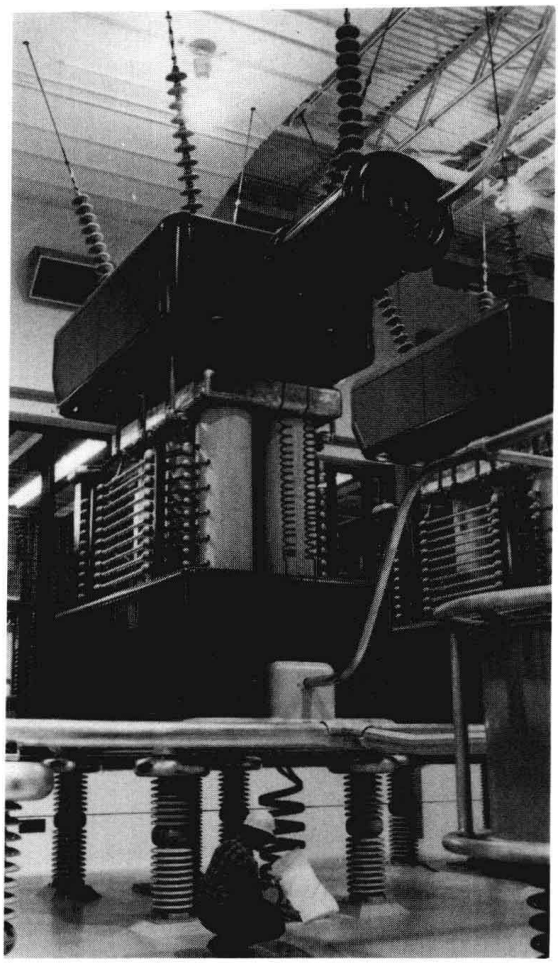
Edison was positive that he was right when he chose direct current as the mode of power distribution for his electrical power installations. Both the voltage and the current were constant at all times, so naturally the amount of power transmitted, their product, was also constant and unvarying. Since the current was unchanging, the magnetic field around the conductors caused by its passage had no appreciable effect on the current's value. Its principal disadvantage was that DC could not be stepped up or down by the use of a transformer as could the "Westinghouse current" (Edison's term for AC). Today engineers with advanced electronic technology at their disposal can generate electrical power as AC and transmit it as DC.

Principles of DC Transmission

The complication involved in converting AC to DC, transmitting it over a high-voltage line, and then converting the DC back to AC can be justified only for transmission lines longer than about 400 to 600 miles. For shorter distances conventional AC transmission lines are still preferred.

The sequence of events in DC transmission is as follows. First, the conventionally generated AC voltage is stepped up by a specially designed power transformer to the required AC secondary voltage (at least several hundred thousand volts). Then, by a process of electronic switching using specially designed mercury-vapor triode tubes (see Chapter 7), the AC voltage is converted to DC.

When an AC source is converted to a DC source by the use of an active device such as a vacuum tube or a solid-state unit, the process is called **rectification,** and the tube is called a **grid-controlled rectifier.** Such tubes have been used in industrial applications for nearly fifty years. Public utility engineers prefer to call such devices **mercury arc valves,** also a good descriptive term. The mercury enclosed in the valve first vaporizes because of the heat and then ionizes, or breaks down into positive ions and



The mercury arc valve (center) converts AC to pulsating DC. The iron-core reactor reduces the intensity of pulses in the output circuit so steady DC results. (Photo courtesy of Bonneville Power Administration, USDOE)

free electrons. The free electrons released in this manner greatly reduce the internal resistance of the tube so that it can carry very heavy currents.

The output of each individual tube in a rectifier installation (actually, six valves are used together) is pulsating (varying) DC. Just as in the power supply of a stereo amplifier, an iron-core reactor (filter choke) is used to smooth the output to steady DC. It is now ready for connection to the conductor of the high-voltage DC transmission line for its long cross-country journey.

Converting DC Back to AC

Assuming that several hundred megawatts of high-voltage DC power have been safely delivered over the transmission line to some distant terminal 700 or 800 miles away, a new problem presents itself: There is no commercial application for DC electrical power at a potential of perhaps 800,000 volts. It must somehow be converted back to AC of the right frequency and then stepped down to a voltage compatible with the terminal city's electrical distribution system.

Michael Faraday discovered in the 1830s that steady DC applied to the primary of a power transformer would not produce AC in its secondary winding. However, he soon learned that by pulsing (interrupting at regular intervals) the DC flowing through the primary, he could produce AC in the secondary. He also found that each pulse of DC in the primary resulted in one alternation, or one cycle, of AC in the secondary.

Public utility engineers employ Faraday's discovery to convert high-voltage DC to high-voltage AC of the proper frequency. They use the same mercury arc valves to rectify AC into DC at the transmission station and from DC back to AC at the line's termination. A schematic drawing of a typical mercury arc valve used in AC/DC and DC/AC conversions is shown in Figure 14.2. The exact technology used is beyond the scope of this book, but a simplified description follows.

The triode tubes used will permit the flow of electrons from cathode to plate only when an appropriate voltage is first applied to the tube's grid, which is mounted astride the electron path from cathode to plate. At the terminal substation a computer analyzes the AC waveform of the power system of the city being served and provides an exactly timed pulse of DC voltage to the grid of each triode tube. Their output is a series of DC pulses of precisely determined duration. These pulses are then fed to

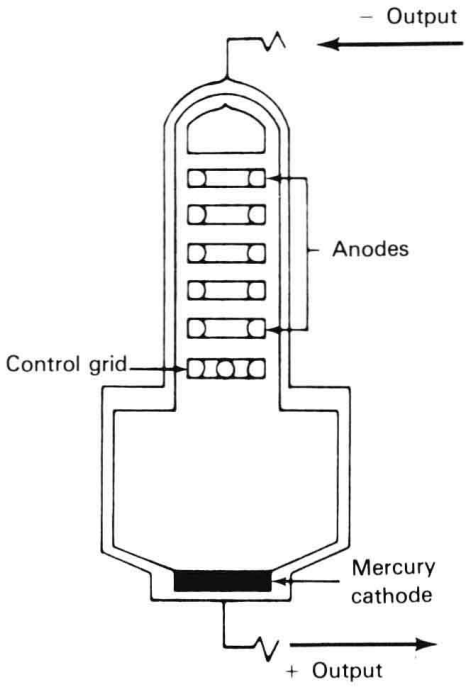


Figure 14.2 Schematic of a Mercury Arc Valve These valves are used to convert AC to DC at the transmission station and to convert DC back to AC at the termination of the power line.

the primary of a very large power transformer capable of handling the entire power output of the transmission line. The transformer's secondary output is AC of precisely the right frequency and exactly the right voltage to be compatible with that of the terminal city's existing power system.

Future Trends

Existing DC transmission systems of the 1980s have proved so efficient compared to their AC counterparts that DC will probably be used on all future transmission lines in excess of 600 miles in length. Existing AC transmission lines covering long distances will probably be eventually rebuilt to employ DC because of the great savings in electrical energy that will result. DC transmission lines are in widespread use in Europe. In the United States, most DC lines are found on the West Coast.

NUCLEAR POWER GENERATION

As far back as 1905, Albert Einstein suggested that an enormous amount of latent power was stored in atoms in the form of electrical bonds linking protons and electrons together in their nuclei as **neutrons.** (A neutron is a paired electron and proton with a net electrical charge of zero.) Einstein reasoned that if this energy could be released in a controlled manner, much of the world's energy demands could be met from this source alone. Much later, scientists suggested that this energy could be recovered by two different approaches: fusion and fission.

Atomic Fusion

On the sun all the heat that serves as the "engine" of the entire solar system is derived by the process known as **fusion.** At the extremely high temperatures of the sun's surface, and under the influence of the tremendously strong magnetic fields existing there, hydrogen atoms with their single proton and single electron are violently torn apart and changed into helium atoms, which have two protons and two neutrons in their nucleus, plus two orbital electrons. *World Book Encyclopedia* estimates that in one second's duration, 657 million tons of hydrogen are changed into 653 million tons of helium on the sun's surface. The discrepancy of 4 million tons represents matter that has been changed into energy—largely radiant energy.

Efforts on Earth to create and control energy by fusion have been largely unsuccessful because our laboratories lack the extremely high temperatures and pressures to make the process a continuing one. Our only application of fusion at present is the hydrogen bomb, which requires an atomic bomb to

trigger it. Once the reaction has been started, nothing can control it, and it lasts for just a few seconds.

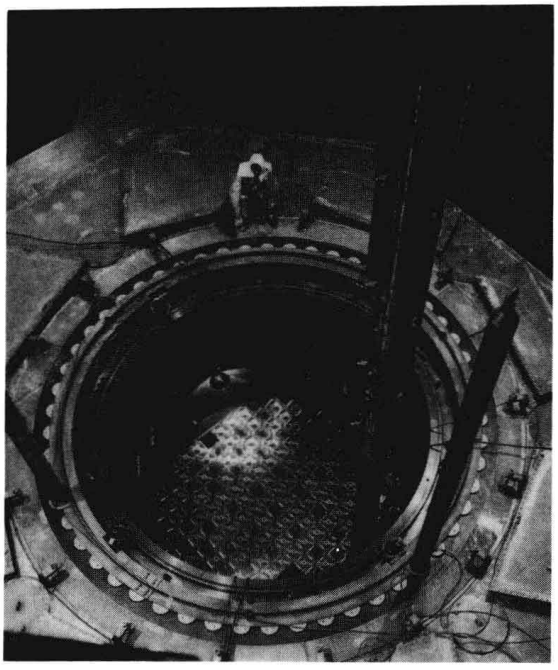
Should we be able to develop and control the fusion process of creating energy on the earth's surface, it would present many advantages. First, hydrogen is one of the most abundant elements on Earth, although most of it is locked up in water molecules. Second, and perhaps of greatest immediate interest, is the fact that most of the by-products of fusion are, to the best of our knowledge, nontoxic. Thus there exists minimum danger that lingering radioactivity will pose immediate physical damage or deferred genetic damage to any form of life. However, scientists agree that the practical development of energy by fusion is at least twenty years away.

Atomic Fission

As we have mentioned, atomic fusion is the process by which hydrogen is converted into helium, a slightly more complex element, and a great deal of heat is given off in the process. With atomic **fission,** we start at the opposite end of the table of 92 elements—that is, with uranium, the most complex of the naturally occurring elements. In fission, uranium naturally decomposes into less complex elements, and again a great deal of heat is given off in the process. By pretreating uranium with neutrons (paired protons and electrons), a more complex form called U-235 can be produced. It disintegrates more rapidly and gives off more heat in a given length of time. It is widely used in the cores of reactors in nuclear power plants.

Nuclear Power Plants

In every atomic plant, heat generated within the core by fission is absorbed by a grid of pipes through which pure water is being circulated by pumps. The water is heated sufficiently to produce high-pressure steam. The steam drives the shaft of a conventional AC generator (or alternator). The steam and water vapor are then condensed in a cooling tower and the water used over again to create additional steam.



This well contains the reactor core of the Trojan Nuclear Plant at St. Helens, Oregon. Water covers the core material. A heat exchanger made of a grid of metal pipes through which water circulates absorbs the heat of the reactor and converts it into steam for turbine operation. (Photo courtesy of Portland General Electric)

In the process of fission, a number of radioactive by-products, such as plutonium, are formed. Extreme care must be taken to see that none of this radioactive material escapes or is allowed to endanger the health of plant employees. Another problem is safe disposal of the spent material of the reactor core. Usually, one-third of the core is replaced every year over a three-year period. At many atomic plants the spent material must be stored under water to minimize the danger of radiation until it can be shipped to a central storage plant eventually to be provided by the federal government.

Although the nuclear fuel of a fission plant could be more effective if it contained a larger percentage of U-235 and an alternate form of uranium, U-238,

the total plant generating efficiency is still very high. The Trojan Nuclear Plant, located on the Columbia River near St. Helens, Oregon, and completed in the late 1970s, can generate up to 1130 megawatts of electrical power. By comparison, Bonneville Dam, located across the Columbia River east of Portland, can produce under conditions of average stream flow only about 1062 megawatts. At present, Trojan is the largest nuclear plant in service in the United States, and the second or third largest in the world.

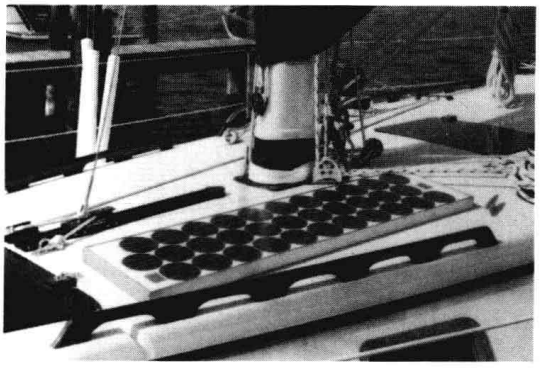
Despite apparent advantages, not a single nuclear power project has been launched since 1978, and work has been suspended on more than a dozen other plants in various degrees of completion. The reasons are many: the failure of the federal government to provide suitable storage facilities for spent core materials; the success of efforts to reduce the nation's total energy consumption; the public apprehension created by the Three Mile Island incident; the constant litigation against nuclear power plant siting, construction, and operation carried on by conservation groups; and the 1983 Supreme Court decision that states have the power to bar construction of future nuclear plants until the government makes provision for nuclear waste disposal. Because of the cumulative impact of all these factors, nuclear power is in a holding pattern for some years to come. In contrast, Western Europe, particularly France, has scores of nuclear plants in operation.

ALTERNATE SOURCES OF ELECTRICAL ENERGY

Most prime hydroelectric sites have already been developed. Yet the supply of fossil fuels continues to wane, and the demand for electrical energy continues to grow at an accelerated pace. Researchers are therefore investigating new energy sources, particularly ones that can be used exclusively for electrical power generation. The federal government has created a Department of Energy to oversee such developments. Its executive officer reports directly to the president. In this section we describe the two most frequently mentioned alternate energy sources—**solar energy** and **wind power**—and appraise their impact on the current shortages of electrical power.

Solar Energy

Solar energy is created by atomic fusion on the surface of the sun as hydrogen atoms burn into helium atoms and liberate enormous amounts of radiant energy in the form of heat and light. Transducers that can directly convert the energy of sunlight into electricity are known as **solar cells,** although the correct term is actually *photovoltaic cells*. (*Photos* is the Greek word for light.) Typically, their output is very low for a single silicon cell—perhaps 0.4 volts at 15 milliamperes. To increase the amount of current available, a large number of cells must be connected in parallel. To increase the amount of voltage available, cells must be connected in series. To increase both the voltage and amperage, the cells are connected in series-parallel, the arrangement used in satellites where large panels of cells are used to charge a low-voltage battery that powers all the communications equipment and sensors aboard.



Heathkit's home electric power kit is installed on shipboard, and the solar panel's bank of photovoltaic cells provides up to 2.6 amperes at 12 volts DC—enough energy to charge a 12 volt battery or operate a small fluorescent lamp. (Photo courtesy of the Heath Company)

Since the cost of enough high-quality silicon solar cells to power a single 50 watt transmitter would exceed $25,000, large installations are impractical until the unit cost per cell can be greatly reduced. NASA is studying two approaches. One is to dispense with the current procedure of first forming the silicon into crystals and then laboriously sawing them into thin, circular sheets with a diamond saw. Direct fabrication from pure silicon promises to be much more economical and equally effective. (In August 1979 the Stanford Research Institute announced development of a new, one-step process for the production of very pure silicon for solar cells. The new technology would reduce the cost of silicon from its present $25.00 to about $2.30 per pound.) A second possibility is to substitute gallium arsenide cells for silicon cells with increased efficiency. Gallium arsenide, already used in the manufacture of LEDs, can withstand much higher temperatures than can silicon and would permit sunlight to be focused into beams of high intensity. Solar panels could then be much smaller but have the same power output.

A number of government agencies, including NASA, are trying to reduce the cost of the manufacturing process (and the resultant cell cost) and to increase the efficiency of the individual solar cells. Under investigation are improvement of the different transducers presently employed, including the original PN junctions (sliced from a silicon crystal); silicon films (both polycrystalline and amorphous); and nonsilicon films such as gallium arsenide. The Department of Energy concedes that although costs have fallen considerably, average efficiencies of solar cells have remained in the range 10% to 25% (the amorphous type averages 10% or less)—too low to be truly cost effective.

Wind Power

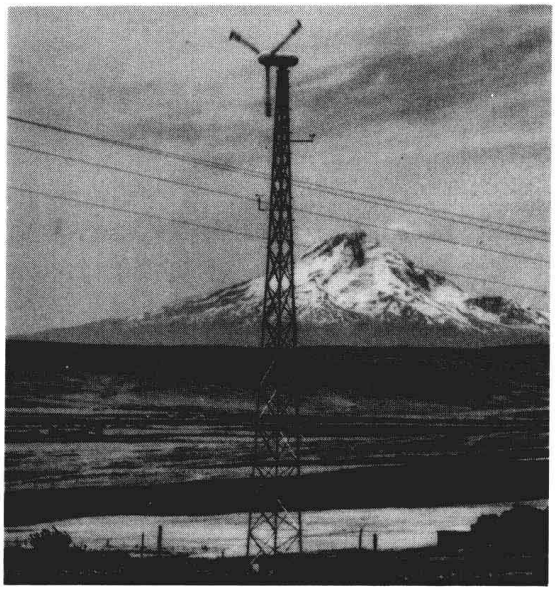
Long before the Rural Electrification Administration (REA) installed thousands of miles of power lines throughout rural America, tens of thousands of farm-lighting plants were in service in the midwestern states. The power source was usually a 32 volt DC generator mounted on a windmill tower. Its shaft was spun at a good clip by a two-bladed propeller facing into the prevailing winds. Its DC output was used to keep charged a very large 32 volt lead-acid storage battery. The power available from a farm-lighting plant was limited. As the name implies, the plants were largely used for lighting homes and barns. With the coming of REA and its readily available 110 and 220 volt AC power, individual farm-lighting plants rapidly fell into disuse. However, the principle on which they operated is still sound today, and individual home installations of the same kind could handily supplement scarce and high-cost electrical power provided by fossil-fuel or hydroelectric plants.

Experimental Installations

In the spring of 1980, NASA, the Department of Energy, and its Pacific Northwest affiliate, the Bonneville Power Administration (BPA), launched two ambitious projects to determine the practicality of wind-generated electrical power. One phase of the project was concerned with low-power generators designed to provide supplementary power for single-family homes; the other, with the generation of large amounts of electrical power to supplement the supply of an existing public utility district and the Northwest power grid.

The siting of these projects was determined only after years of in-the-field surveys by staff members of Oregon State University. The location finally chosen was Goodnoe Hills near Goldendale, Washington. Goldendale is located near the channel of the Columbia River just to the east of where it breaches the Cascade Mountains. This topography strengthens the prevailing west-to-east winds to create one of the highest average wind velocities in the Pacific Northwest. The electrical output of a wind-driven generator varies as the cube of the wind speed. Therefore, even a small increase in wind speed markedly increases the power output.

Nine comparatively low-power machines were selected by BPA for their supplementary power project. The machines ranged from 2 to 4 kilowatts in capacity and were designed to provide an output



BPA's small-scale wind turbine spins as an easterly wind blows down the Columbia River. Each turbine supplies 2 or 4 kilowatts of electrical power to a single residence. (Photo courtesy of Bonneville Power Administration, USDOE)

voltage of 110 volts at 60 hertz. Mounted on 60 foot self-supporting towers, each power unit had its input fed to a nearby home through a low-voltage underground cable. Since output was 110 volts, 60 hertz, like the existing public utility district power, there were no problems in the interconnection of the two power sources. However, the wind-generated power could only supplement the household power needs by supplying perhaps 40% of the total electrical requirement. No serious electrical or mechanical problems developed, and this project is still in operation.

Underwritten by the Department of Energy and NASA, the second project had an entirely different goal—creation of the world's first wind farm to generate as much electrical power as possible for sale to northwestern public utilities through the distribution facilities of the Bonneville Power Administration. Three very large wind turbines were constructed for this project. The propeller-like rotors were 300 feet long, and the power-generating unit was mounted in a streamlined nacelle 200 feet above ground level. (The Boeing Company designed and constructed the wind turbines.)

Each wind turbine is capable of generating 2.5 megawatts of power at wind velocities between 27 and 45 mph and lesser amounts at lower wind speeds. At speeds higher than 45 mph, the blades automat-

BPA's large wind turbine has a turning blade 300 feet long. Its computer-controlled power-generating unit is 200 feet above the ground. Power output is 2.5 megawatts in a wind of 27 to 45 mph. Higher velocities cause the computer to feather the rotor. (Photo courtesy of Bonneville Power Administration, USDOE)

ically feather to protect the generating equipment from mechanical and electrical stress. Feathering cuts off all electrical power output. (On-board computers are in full control.)

The use of a few large wind turbines is being compared to the use of large numbers of low-output generators. The supersize turbines, which cost $5 to $6 million each, have undergone mechanical stresses in adverse wind conditions that forced their shutdown for repair or, in some cases, their redesign and rebuilding. Preliminary information seems to indicate that future installations should favor wind machines of lesser mechanical size and smaller electrical output. The use of more and smaller machines could produce an equivalent amount of power with far fewer mechanical breakdowns.

Metglas and AC Power Distribution

Closely related to alternate energy sources is the development of a glasslike, iron-based alloy that promises greatly to increase the amount of energy available from existing hydroelectric and fossil-fuel generating plants. This alloy will substantially reduce the core losses of the hundreds of thousands of distribution and step-down transformers used by all electrical power systems. The new core material, called *Metglas*, is amorphous—that is, lacking crystalline structure. Because of its structure, much less electrical energy is wasted in magnetizing and demagnetizing the core each time the cycle changes, and energy saved is additional energy that can be used by consumers.

SUMMARY

Since Thomas Edison had found that electrical power distribution at low voltages and high currents (DC) was practical only in densely populated urban districts, its distribution to suburban and rural areas seemed in doubt. This problem could not be resolved until George Westinghouse introduced the AC system of power generation and distribution, which permitted whole cities to be served by a single system with minimal losses. The Westinghouse system was based on efficient AC generators (alternators) linked to large, efficient power transformers capable of stepping the AC from the generators up to high voltages (and low currents) for long distance transmission. Then additional transformers were used to step down the voltage and step up the current for local distribution and consumption.

Nearly a century later the same concepts are still used in AC power generation and distribution. Only after about 1960 was it realized that the limits to AC distribution (for technical reasons) were being approached, and engineers turned their skills toward the design of transmission systems based on extremely high DC voltages generated by sophisticated electronic circuits. More power could thus be transported with lower electrical losses.

Well over 90% of the world's power is generated electromagnetically by revolving at high speeds coils of wire wound over an iron core (armature) in the intense magnetic field produced by heavy DC currents passing through stationary iron-core coils (field coils). All generators produce AC currents internally, but some are equipped with commutators and brushes to change the AC current into pulsating DC (DC generators). Others, using brushes and slip-rings, lead the AC current unchanged to the output terminals and are called alternators.

Turning the armatures in the intense magnetic field requires very large amounts of power, which comes from two principal sources: hydroelectric and high-pressure steam. Most steam plants burn increasingly scarce and expensive fossil fuels. Others use the heat generated by a nuclear reactor as a heat source, but the number of nuclear plants is still small, and their future expansion is in question.

The theory behind the use of existing AC transmission systems is that the comparatively low AC voltage and high current output of the generators is transformed by a very large transformer to a very high voltage and a very low current. This procedure makes the current losses due to the small resistance of the conductors of the high-voltage line completely

acceptable. However, when the AC line voltage was increased to around 800,000 volts to permit even larger amounts of power to be transmitted, problems arose with insulators and with the spacing of conductors.

Meanwhile, electronics was providing engineers with new approaches that made DC transmission feasible and attractive: high-voltage rectifiers that could be used first to convert AC to DC at the generating site and then used in a switching circuit at the termination of the line to convert the DC to pulsating DC. Another contribution of electronics, a highly specialized computer, times and supervises this operation with split-second accuracy. When pulsating DC is applied to the primary of a substation's step-down transformer, the frequency of the AC appearing in its secondary windings is precisely the same as that flowing in the distribution system of the terminal city.

With the rapid increase in the cost of fossil fuels (particularly petroleum) as demand exceeds supply, and with the safety of nuclear plants under question, engineers are seeking alternative sources to replace steam generating plants. (Most hydro sites have already been developed.) Solar energy is already being used for space heating and the heating of water, but is an inefficient direct means of producing electrical energy. Present solar cells are only about 20% efficient. Wind power presents even more problems: the lack of suitable sites near population centers, the inefficiency of converting mechanical motion of propellers electromagnetically into an AC or DC current, and the inability to interconnect wind generator plants with each other or with conventional power systems. At this time, wind farms seem to be more practical than acres of solar cells.

KEY WORDS AND CONCEPTS

alternator
dynamo
fission
fossil fuels
fusion
generator
grid-controlled rectifier
hydroelectric power
mercury arc valves
neutron
rectification
solar cells
solar energy
telemetering
transmission line
wind power

SELF-TEST QUESTIONS

1. How did Edison and Westinghouse differ in their approach to providing electrical power to a large metropolitan area?
2. What are the two major sources of mechanical power for the nation's electrical generating plants?
3. How did the development of transmission line technology make possible the utilization of our country's finest hydroelectric sites?
4. What is the advantage of generating AC power at a low voltage and a high current and transmitting it at a high voltage and a low current?
5. Describe the sequence of events in high-voltage DC power transmission from its original generation as AC to its terminal delivery as DC.
6. When electrical power is transmitted over great distances as a steady DC current, why is it nec-

essary to convert it to pulsating DC at its termination before applying it to the primary of the substation transformer?

7. Why is it true that all prime sources of electrical power (sunlight, wind, falling water, wave action, fossil fuels) are directly or indirectly linked to solar energy?
8. Explain why the alternators in a nuclear power plant are driven by steam turbines.
9. Why is it uneconomical to generate large amounts of electricity by the use of solar cells?
10. In wind-driven generators used in farm-lighting plants, why was all or part of the generator's output used to charge a storage battery of very large capacity?

15 Electronics Today in the Home, Business, and Industry

One way for a family to enjoy its color TV receiver is by connecting the antenna terminals to a video disc player, which permits viewing of favorite programs at any convenient time. Visual and aural quality are excellent. (Photo courtesy of RCA)

Electronics in the Home
- Microwave Ovens
- Solid-State Laundry Units
- Smoke Detectors
- Photography and Cameras
- Electronic Weather Stations
- Color Projection Television
- Video Games
- Cordless Telephones

Electronics in Business and Industry
- Telephone Answering Systems
- Electrostatic Copying
- Word Processing
- Robots and Robotics
- Optoelectronics

Latest Trends in Electronics

Summary

Key Words and Concepts

Self-Test Questions

In the first section of this chapter, we discuss new uses of electronics in the home. We begin with new household appliances that use electronics as an aid to operation and then move to hobby and entertainment-oriented applications—for example, cameras, weather stations, video games, and cordless telephones.

The second section of the chapter focuses on applications in business and industry. It opens with a discussion of the Electronic Industry's most rapidly growing sector, robotics. We then turn to more conventional industrial applications. Our final topic is a technology that may revolutionize telecommunications: optoelectronics, the transmission of data from a varying light source along an optical fiber as a conducting medium.

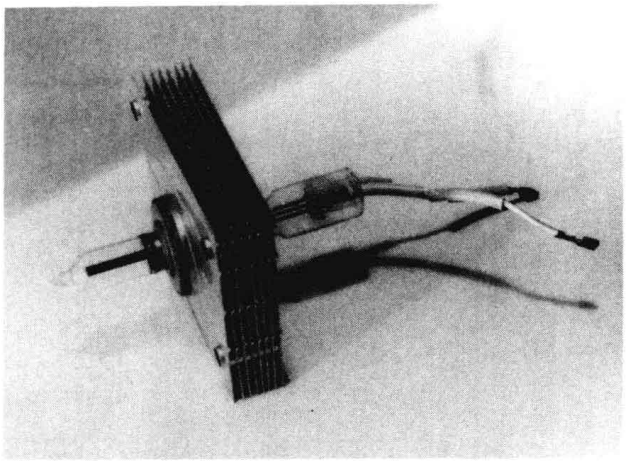


Figure 15.1 Raytheon Magnetron Tube Aluminum fins provide cooling to the tube's structure, which is mostly to the right. The part to the left of the fins is the tube's radiating section, which feeds energy into the oven's waveguide for distribution.

ELECTRONICS IN THE HOME

Microwave Ovens

Sales of microwave ovens have shown a slow but steady increase over the past five years, and appliance dealers consider them one of the most promising sources of future income, particularly since color television sales have apparently leveled off or even shown a decline.

Microwave cooking is based on the law of conservation of energy: Energy can neither be created nor destroyed, but its form can be changed. In a microwave oven, two changes take place in the form of energy. First, 60 hertz energy from the power lines is converted into 2450 megahertz energy by the action of a **magnetron** vacuum tube in an oscillatory circuit involving a strong electromagnetic field. Figure 15.1 shows a Raytheon model 1381-B magnetron. It is about 5 1/2 inches overall and is cooled by aluminum fins that conduct heat away from its anode, or plate.

When the microwave energy generated by the tube is conducted to the food to be cooked, the food's molecules are set violently in motion. This molecular agitation and friction produces the heating and cooking. Microwaves travel freely through conventional glass and porcelain dishes, which, like the interior of the oven, remain cool throughout the cooking cycle. In general, the greater the moisture content of the food, the greater the amount of heating produced by the microwave radiation and the faster the cooking process.

The electronic technology involved in the microwave oven is shown in Figure 15.2. First, the microwave energy leaves the end of the magnetron tube and travels down a metallic tube or duct called a **waveguide.** There it encounters the metallic blades of a slow-turning electric fan known as a **stirrer.** The function of the stirrer is not to provide cooling or ventilation but to uniformly distribute microwave energy over the entire surface of the food being cooked. The stirrer distributes energy by reflection of the microwaves from its revolving metallic blades. The **interlocks** on the left are merely series switches installed in the 110 volt AC line, which supplies power to the magnetron. Their use makes it impossible to operate the microwave oven without the oven

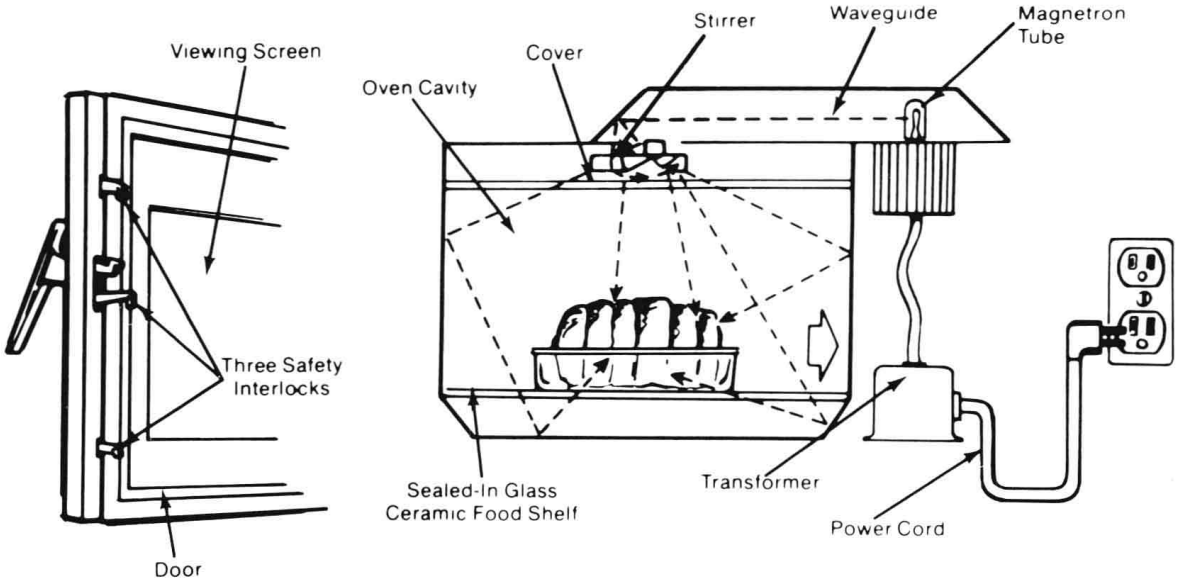


Figure 15.2 Whirlpool Microwave Oven The diagram shows the power supply, magnetron tube, waveguide, stirrer, and oven cavity. (Photo courtesy of Whirlpool Corporation)

door being closed and securely latched—an important safety factor.

Recent models of microwave ovens by all major manufacturers include in their deluxe line one or more microprocessors as control devices to supervise the oven's operation. **Microprocessors** (see Chapter 13) are essentially a very elemental form of microcomputer, though they lack a microcomputer's sophistication, particularly memory. They are usually built around a single integrated circuit engineered for this purpose and are supplied by firms such as Intel.

Commercial microwave ovens vary widely in power consumption, ranging from as little as 400 watts in economy models up to about 1500 watts in larger models. In the mid-1980s the electronics of microwave ovens has increased in sophistication. The microprocessor unit has both analog output (for control) and digital output. The digital output is used to drive through a converter circuit the digital clock's readout. Also, the microprocessor has more programmable memory available through the push buttons on the front panel.

Solid-State Laundry Units

The term *solid state* is a little misleading as applied to washing machines intended for home laundry use since the size of the electronic control device is very small in proportion to the size of the appliance itself. Nevertheless, the control device works very effectively. Whirlpool was the first major appliance manufacturer to ship so-called solid-state washers to its dealers, and by the late summer of 1978 it was shipping microprocessor-controlled washer-dryer combinations as well.

Electronic control devices are needed because of the popularity of synthetic fibers and permanent press garments, which are sensitive to the temperature of the water in the wash cycle, the length of the cycle, and the temperature of air and the length of drying cycle. Through the use of electronics applied in the form of a microprocessor that controls both washing and drying cycles, water temperatures and length of the cycles are always at their best levels. Once the buttons on the control panel are pressed, the microprocessor takes over, and no fur-

ther attention is required until the laundry is completely dried. Like microwave ovens, "smart" washers and dryers of the mid-1980s employ more sophisticated electronics than appliances of the late 1970s. An additional advantage of solid-state laundry units is that the washer can be programmed to use cold water, which saves energy.

Smoke Detectors

Some 300,000 Americans are injured by fire each year. Of this number, 12,000 will die; another 50,000 will be hospitalized for six weeks to two years at an average cost of $8000 per patient. Residential fires account for most fire deaths, and most of these deaths take place at night. Fire prevention officials say that injuries and deaths could be sharply reduced if every home were equipped with at least one smoke detector that would sound an alarm as soon as a fire started. Millions of smoke detectors have already been installed. By the mid-1980s many cities required their installation in public housing.

Two common types of smoke detectors are offered for sale: the **ionization chamber** type, which is comparatively new but is the most widely distributed; and the **photoelectric cell** type, which has been widely used in industrial security systems for many decades but has recently been adapted to home use. Ionization chamber types can be manufactured more inexpensively. An interior view of a Honeywell smoke alarm unit of the ionization chamber type is shown in Figure 15.3.

The key item in an ionization chamber detector is a tiny amount of the radioactive element americium, produced by the bombardment of uranium by high-energy helium nuclei. A grain or two of the material, wrapped in gold foil for protection, is sufficient. Mounted in a sensing chamber, it ionizes the surrounding air to make it conductive. (When a substance is ionized, its atoms lose one of their valence electrons to produce ions with a positive charge and free electrons with a negative charge.)

Powered by a 9 volt transistor battery of the heavy-duty alkaline type, a flow of electric charges takes place across an air gap to which the battery is connected. If fire-caused ions in the form of gases or smoke enter this ion chamber, they interfere with the movement of electric charges across the air gap, and the battery current falls. The change in current, sensed by appropriate electronic circuitry, triggers the alarm. The alarm is basically an audio oscillator tuned to a frequency of several hundred hertz, which in turn actuates a very efficient and compact loudspeaker. Its sensitivity can be varied over considerable limits by adjusting a variable resistor. Too much sensitivity can cause frequent false alarms; too little can delay a warning of a real fire.

Figure 15.3 Interior View of a Smoke Alarm This Honeywell unit is an ionization chamber type. The ion chamber is to the lower left. Sensitivity adjustment is to the upper left. Audio alarm (actually a loudspeaker) is to the lower right, below a CP-7737 integrated circuit.

Photography and Cameras

At least 50 million amateur photographers have purchased Kodak's 110 series pocket cameras since their introduction in 1972, and sales have been running at the level of at least 5 million additional units a year. Today the 110 format is used by at least seventy manufacturers of similar automated cameras, both American and foreign. Certain models may use elec-

tronics in at least two ways to automate their performance.

While some models still use two conventional 1 1/2 volt cells in series to fire the filament material in the flash cube, others may use a piezoelectric transducer for the same purpose. When the shutter is operated, a crystal of barium titanate (or similar material) is momentarily stressed to supply the firing voltage.

A cadmium sulfide (CdS) light cell with its own window on the front panel is frequently used to control the shutter speed automatically under varying light conditions. The cell behaves as a **varistor**—that is, its resistance varies with light intensity. It is connected in series with a tiny D'Arsonval galvanometer type of motor that mechanically moves the shutter speed adjustment. In turn, the shutter adjustment is connected in series with a special voltage source using a dime-sized cell of the silver or mercuric oxide types having a potential of about 1 volt.

When the light is very bright, the resistance of the CdS cell drops, the motor armature turns against the tension of a spring, and the camera's shutter is set for its fastest exposure. Conversely, when the light is dim, the resistance of the CdS increases, the motor develops less turning motion, and the shutter is set for the slowest exposure. When the light is of some intermediate value, the shutter speed is set to a medium level.

Varying shutter speeds over wide extremes may pose several minor problems. First, when the shutter speed is around 1/25th of a second, it becomes difficult to hold the camera steady enough to produce a sharp picture. At the same time, any action within the photo is almost sure to be blurred. No focusing problem occurs since the camera has a fixed-focus lens that will produce a picture reasonably in focus from three or four feet to infinity. Sharper, more professional pictures require a single-lens reflex (SLR) camera, which costs several hundred dollars more.

Single-Lens Reflex Cameras

SLR cameras use more elaborate electronics than the 110 series. Newer models frequently use silicon photocells rather than CdS types. Silicon photocells are faster-acting and produce their own voltage—about 1/2 volt or less. Mounted within the camera case, they sample the light as it is brought to focus by the lens—a much more accurate way to sample light levels because the cells' output voltage is directly proportional to the light intensity. The output of the silicon cell consequently can be used in a number of ways to ensure correct film exposure. In the simplest and least expensive application, it is used to operate a D'Arsonval meter, which can be seen in the camera's viewfinder. The operator merely reads the f-stop reading indicated by the meter and then manually sets the aperture adjustment on the lens barrel to correspond.

Another use of the silicon cell's output is to control automatically the camera's shutter speed. The same galvanometer movement can be used to control the aperture opening rather than the shutter speed or to control both shutter and aperture automatically (an option that greatly increases camera cost).

Light sources for the SLR and similar cameras are making a rapid transition from the filament flash bulb ignited by a low-voltage battery to the **strobe flash tube.** Electronically, strobe flashes are quite simple in design. A transistor is connected in an oscillator circuit operating in the ultrasonic range and generates a voltage of at least several hundred volts but at a very low value of current. This voltage is used to charge a capacitor. A mechanical switch in the camera's shutter linkage operates a trigger circuit in the light unit, which applies the charged capacitor across an argon-filled strobe lamp. Usually, the camera's shutter speed must be reduced to 1/60 or 1/30 second to maintain flash synchronism. The flash unit's basic power requirements are met by the use of two heavy-duty alkaline cells. In more expensive flash units a fast-acting silicon cell measures the light bounced back from the subject and adjusts the flash's intensity to the proper value. The actual exposure time is basically the length of the flash, not the camera shutter setting.

New Camera Models

Late in 1979 Polaroid announced production of their Sonar camera (Figure 15.4), which not only

Figure 15.4 Polaroid's Sonar Camera This SLR camera features autofocus and electronic flash. A panel-mounted transducer beams an ultrasonic pulse to the subject and also picks up its reflection. A microprocessor then computes the exact distance and controls the motor that brings the lenses into focus. (Photo courtesy of Polaroid Corporation)

adjusts the shutter and aperture for proper exposure but also measures the distance between lens and subject and then sets the lens to the point of sharpest focus. The camera sends out an ultrasonic wave traveling at 1100 feet per second. The almost instantaneous echo is picked up by a microphone and the elapsed time measured by an integrated circuit. A microprocessor rotates the lens assembly by a predetermined amount to achieve perfect focus. Since this process is based on ultrasound and not visible light, the focusing can be perfectly done in complete darkness.

The 1983 models of both Pentax and Canon achieve self-focusing through a rather complex system that analyzes, with an array of photodiodes, the relative contrast of various areas of the image. (Maximum contrast means the image is in focus.) A microprocessor evaluates this data and controls the movement of the DC motor that moves the lens barrel until maximum contrast is reached.

Electronic Weather Stations

Sales of electronic equipment for home weather stations have increased dramatically. In the typical weather station kit, only the barometer is nonelectronic. A conventional aneroid type is used. Of the remaining instruments, the simplest is the wind-direction indicator. The sensing device itself is mounted on a rooftop mast. As the wind vane turns to face into the prevailing wind, its shaft turns a switch rotor across a piece of insulating material on which are mounted eight contact terminals, each of which is wired to a lamp behind the compass points on the instrument's dial. When the circuit is closed, the appropriate lamp lights.

The wind-speed indicator, or **anemometer,** is a form of AC generator in which the whirling cups are attached to the same shaft as a rotating armature, which revolves in the field of a compact permanent magnet. The wires carry the AC voltage generated down the mast and to the instrument on the control panel. The D'Arsonval meter actually measures the amount of voltage generated by the rotating armature, but its scale is calibrated in terms of miles per hour. A series resistor of the appropriate value, by switching, changes the maximum wind velocity that can be measured from 30 to 90 mph. A diode is used to change the AC to DC for proper operation of the meter.

Indoor and outdoor temperatures are continuously sampled by temperature sensors, which are actually transducers. Their output is a DC voltage directly proportional to the temperature applied to the sensing device. Accordingly, the meter used is actually a very sensitive DC voltmeter calibrated in degrees Fahrenheit. Other weather panels are available with such features as digital readout using LEDs and scale readings in kilometers per hour and degrees

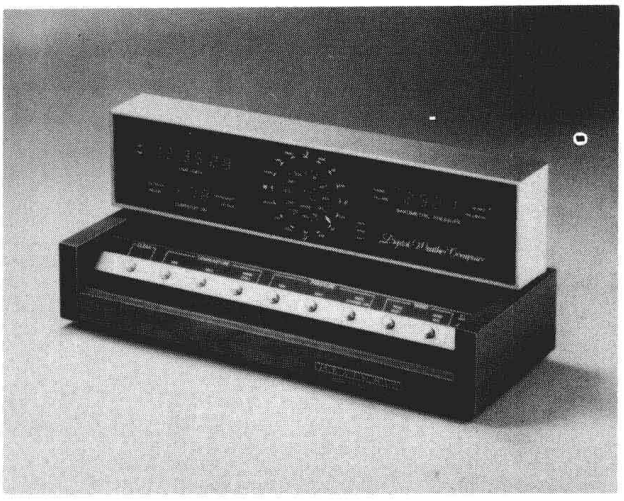


Figure 15.5 Computerized Weather Station Heathkit's model IDW provides weather watchers with data on wind direction and velocity, barometric pressure, and outside temperature. A rooftop anemometer and wind-direction indicator provide data over a multiwire cable. (Photo courtesy of the Heath Company)

Celsius. The unit shown in Figure 15.5 is powered from the 110 volt line through its own power supply.

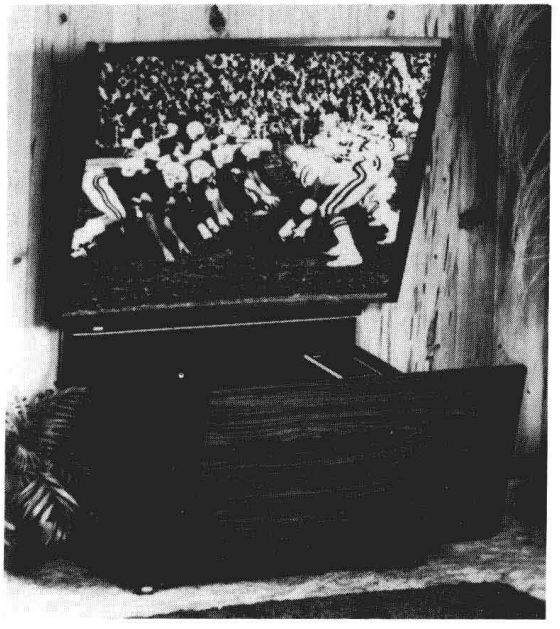
Color Projection Television

In the future, wall-screen television will be commonplace. General Electric, a pioneer in the field of television, has recently produced its Widescreen 1000, which has over three times the screen size of the largest conventional TV set sold today. The screen is 27.4 inches by 36.6 inches and totals more than 1000 square inches in area, yet this increase in viewing area has not sacrificed definition, color quality, or reliability.

The large screen has been created by use of a 13 inch CRT of great accuracy and brilliance whose image is magnified many times by a highly engineered optical system employing an **acrylic lens** and two mirrors. The basic arrangement of units is shown in Figure 15.6. A further refinement in the design is the addition of a viewing screen employing a sheet of acrylic plastic molded in the pattern of a **Fresnel lens,** a technique that results in uniform distribution of light over the 1000 square inches of surface without noticeable hot spots or darkened corners.

By using a 13 inch CRT with a high-voltage requirement of only 32 kilovolts (typical of larger color TV sets), a very large picture can be produced with a standard GE television chassis. The same GE modular solid-state chassis used for their line of 25 inch consoles is used for the Widescreen 1000.

Other manufacturers have used the approach of employing three powerful projection tubes of small diameter but great brilliance (one red, one green, and one blue). These tubes are precisely focused,



RCA's model PPB-100R projection color TV provides a vivid, highly detailed image that fills a 50 inch viewing screen. Remote control is provided by means of a low-power infrared transmitter/receiver combination. (Photo courtesy of RCA)

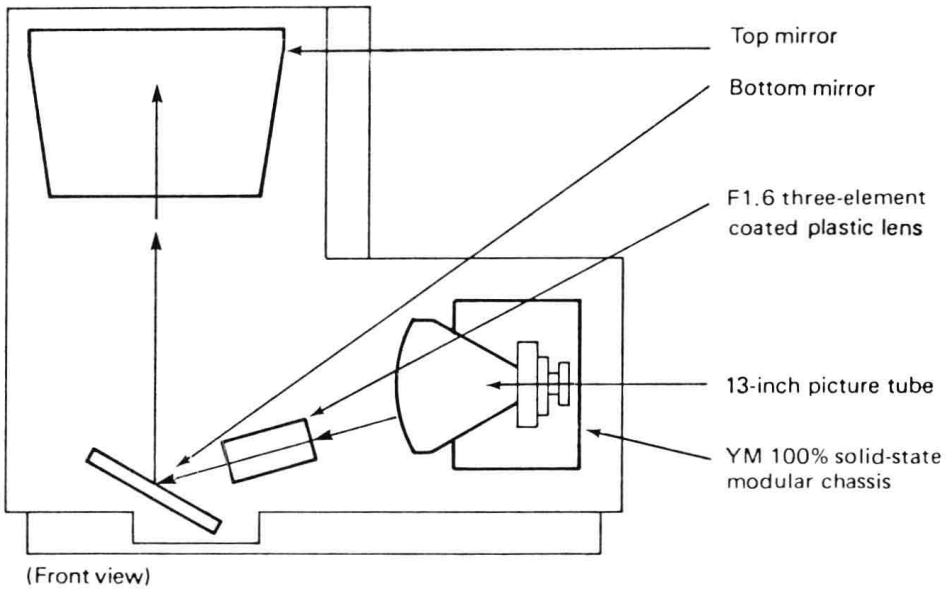


Figure 15.6 Schematic of GE Widescreen 1000's Projection Path Note the use of a plastic lens and two mirrors in the optical system. Not shown is the Fresnel lens (behind viewing screen) used to avoid hot spots and dark corners. (Photo courtesy of General Electric)

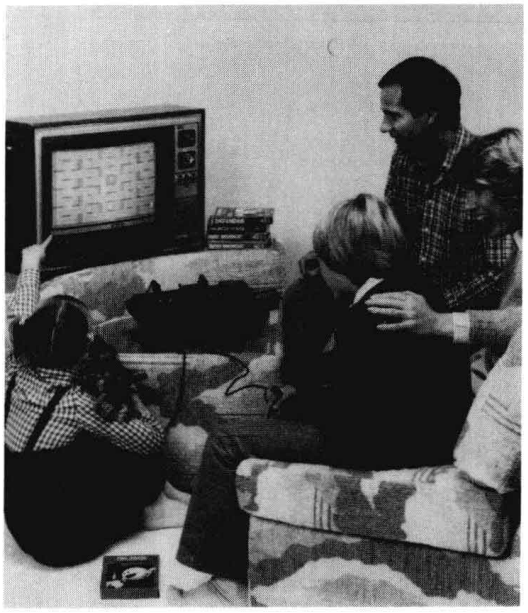
by a system of lens and mirrors, on an external opaque screen similar to that used for home motion picture and slide projection. This approach produces a larger and more brilliant image but is considerably more expensive to manufacture.

Projection TV sales, now quite slow, are expected to show steady growth during the next five years. Production of a flat wall screen less than an inch in thickness will come about only when we learn to diffuse and use liquid crystals more effectively. Encouraging progress in black-and-white TV is now being made along this line. The closest approach to a liquid crystal display has been Seiko's flat screen color TV announced in the fall of 1983. Less than half the size of a typical business card, the display is housed, sandwich-like, between two thin pieces of glass. Special liquid crystals with faster-than-usual response time are used. The display contains about 60,000 **pixels** (picture elements), each comprised of liquid crystal groupings, a thin-film driver transistor, and color filters. Mitsubishi is also working on production of much larger screens.

Video Games

Video games, once seen only in amusement arcades, are now selling to home users. Despite sharp price cuts in player units and game cartridges, sales are slow because of the many new manufacturers entering this now overcrowded market. Many of the newer personal computers come equipped with a chassis socket that will accept standard plug-in game cartridges.

A video game player is essentially an elementary microcomputer. The microcomputer receives its operating instructions by means of a plug-in game cartridge that provides the read only memory (ROM) required for each game. (The cartridges usually contain one or more LSIC units and nothing else.) The ROM supplies the TV set (through the modulator



A family prepares to play a video game on their Atari model 2600 player. Games are played by plugging in a game cartridge and then moving the two joysticks that control ganged variable resistances or switches. Game players are actually minicomputers with minimal amounts of memory. The color TV set is entirely separate. (Photo courtesy of Atari)

unit) with the necessary stabilizing pulses, picture background, and sound effects. Moving the handles on the game controllers produces a voltage or current change that causes figures to move across the TV screen. Figure 15.7 is a block diagram of a video game's operation.

Unlike the very good visual quality of arcade video games, visual quality of the typical home unit is disappointing. Its picture elements (pixels) are too large because of insufficient capability in the plug-in game cartridge. The only remedy is to increase the amount of memory. Future video games will probably use video disc techniques.

Cordless Telephones

With the new cordless telephones, a call can be taken a quarter-mile or more from the base station in the house. Audio quality is equal to, or better than, that provided by a wired telephone.

Every cordless telephone has two basic units. The AC-powered base unit plugs into the home's nearest telephone jack and a nearby AC 110 volt outlet. Its electronics include a low-power AM transmitter operating on an FCC-approved frequency near 1.7 megahertz; a provision for using this transmitter to superimpose the radio frequency signal on the 60

Radio Shack's model ET-330 cordless telephone uses a built-in microprocessor to offer features not available on less complex units: a programmable security code to prevent neighbors from dialing on the line, automatic dialing of two 10 digit numbers, and automatic redialing of a busy number. (Photo courtesy of Radio Shack, a Division of Tandy Corporation)

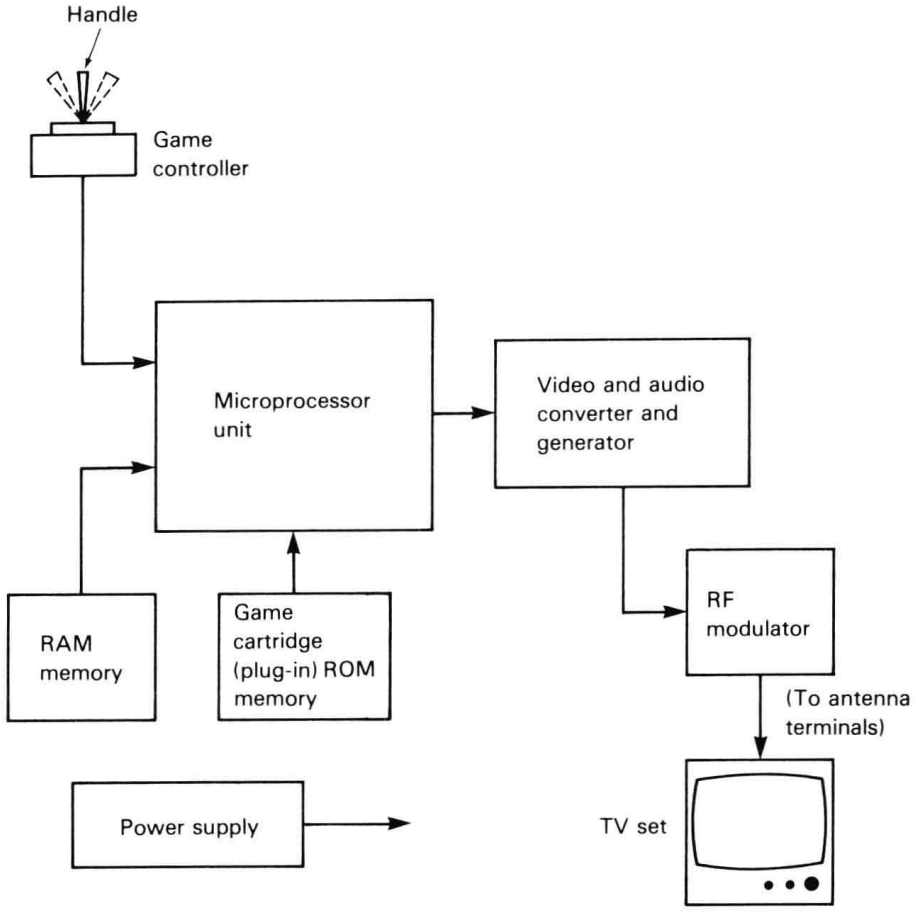


Figure 15.7 Block Diagram of a Video Game A video game player is a simple computer.

hertz line current as a carrier current; a sensitive FM receiver for picking up the transmissions of the handset on about 49 megahertz; and a battery charger to recharge the Ni-Cad batteries that power the handset. The house wiring serves as the antenna for its 1.7 megahertz AM transmitter.

The other basic unit of the cordless telephone is the handset, which is much smaller than on conventional, wired telephones. The handset contains a 49 megahertz narrow-band FM transmitter connected to the telescoping antenna; a sensitive, fixed-tune 1.7 megahertz AM receiver; a microphone and earphone; and a touch-tone pad for dialing. The unit actually uses pulse dialing since it can be used with any existing telephone system. (It may include provision for redialing the last number called and for holding up to two frequently called telephone numbers in memory.) The handset is powered by Ni-Cad batteries that are automatically recharged when the unit is replaced on the cradle of the base unit. Operation is identical to that of a conventional extension telephone.

ELECTRONICS IN BUSINESS AND INDUSTRY

In this section we discuss some familiar electronics applications in business and industry and explain the basic electronics involved. We then focus on two rapidly developing areas: robotics, which promises dramatic changes in the use of labor in manufacturing; and optoelectronics, which is already having impact on the telecommunications industry.

Telephone Answering Systems

Many small businesses are turning to telephone answering systems to minimize interruptions in work routines. Although the electronic principles involved in the operation of such devices are very simple, few people who use such units daily have any idea of how they work.

What follows is a description of what happens from the time an incoming call is received until the caller hangs up. The telephone ring that initiates the chain of events is actually an AC signal of several volts' intensity. This AC voltage is rectified to DC by a solid-state diode that energizes a relay to start the AC-powered tape recorder with its prerecorded message.

Radio Shack's model TAD-112 telephone answering machine uses two audio cassettes for maximum flexibility in recording incoming calls and delivering prerecorded messages. It can record up to 120 calls and screen incoming calls. The owner can also dial the instrument from elsewhere and have it play the recorded messages. (Photo courtesy of Radio Shack, a Division of Tandy Corporation)

Recording of incoming telephone messages comes under FCC jurisdiction. Current regulations do not require the use of a regular beeper tone while the recording is under way. The message of intent to record and the ensuing audio tone suffice. In certain circumstances, such as when required by state law or by local telephone company rules, an approved type of electronic interface must be used to isolate the answering unit from the telephone itself.

As soon as the tape recorder starts recording the incoming call, a solid-state diode is sampling part of its output amplifier voltage and rectifying this voice signal into DC. This DC is then used to keep the *on* relay in the *operate* position. As soon as the person stops talking, this rectified voice signal disappears, the relay drops out, and the answering unit is again in the *ready* state to await the next call. Playing back the recorded messages is as simple as operating any cassette tape recorder.

The telephone answering system can also be used to screen incoming calls. The monitoring is done by means of a small loudspeaker connected across the recording unit's audio amplifier. The owner can decide whether to bypass the system and carry on a normal telephone conversation. To do so, it is necessary only to push a button.

Electrostatic Copying

Thirty years ago businesses had just one way to make multiple copies of a letter—by the use of carbon paper. Four or five carbon copies could be made at the same time, but seldom did the carbons match the original in clarity and neatness. A typing error was a real calamity: A mistake in the original and in each carbon copy had to be separately erased and retyped.

Some experimentation was taking place with photographic processes, but they proved too expensive for daily use. In 1946 the Haloid Company, a

The Xerox 5400 copier and duplicator allows high-speed duplication; two-sided copying; versatility in handling solids, halftones, and light or dark originals; and collating. A microprocessor supervises the overall operation and remembers where operations ended when the stop button is pressed. (Photo courtesy of Xerox Corporation)

firm offering a complete line of photocopy products, decided to work jointly with the Battelle Memorial Institute to explore the possibilities of a new, dry process of electrophotographic copying that had been developed and patented by Chester F. Carlson. The process became known as **xerography** (a combination of *xeros*, the Greek word for "dry," and *graph*, an abbreviation of the Greek verb "to write"). At this stage xerography was beset with technical problems, but by 1958 the process was close enough to commercial production to warrant the Haloid Company's changing its name to Xerox Corporation.

The process of xerography is based on one of the statements of Coulomb's Law of Charges: Like charges repel, and unlike charges attract. The copying process in its simplest form is based on producing on a sheet of paper an electrostatic charge conforming to the printed or written original and then dusting it with powdered ink particles having the opposite electrical charge.

Putting theory into practice is somewhat more complex. Figure 15.8 shows the sequence of events that occurs. A rotating aluminum drum coated with selenium, a photoelectric material, is exposed to a source of reflected light as it turns. The light is a reflection of the image from the brightly illuminated sheet of paper. As this bright image falls on the selenium-coated cylinder with its positively charged particles, the bright areas lose their positive charge. The dark areas, representing the printed materials of the original page, retain their positive charge.

To bring out the image, it is necessary only to dust the rotating drum with powdered ink whose particles carry a negative charge. They adhere immediately to the positively charged areas, and then a mechanical process transfers the pattern on the inked drum to an ordinary sheet of untreated paper. To lock the ink particles into the surface of the paper, momentary application of heat is used. The still-warm print is often difficult to tell from the original. The process described is used in one of the simpler tabletop machines, the model 914 copier. Much more sophisticated models are now in use, with capabilities of up to 2400 copies an hour and with variable image size reduction. The principle on which they operate is the same as that on which simpler copiers are based. The more advanced 1984 models feature an electronic troubleshooting chart that shows the repairer exactly where to look for a malfunction.

Word Processing

Thirty years ago secretaries were still typing letters, processing data, and storing information in about the same manner and with approximately the same efficiency as secretarial help of the early 1900s. Business and commerce had, in the same period, made vast strides in efficiency, particularly in the day-to-day handling of information and its eventual storage for future use. It became imperative for business to rethink its archaic ways of handling words and the ideas they represent.

Not the ultimate answer, but certainly a step in the right direction, was IBM's magnetic tape Selectric typewriter introduced in 1964. A secretary could

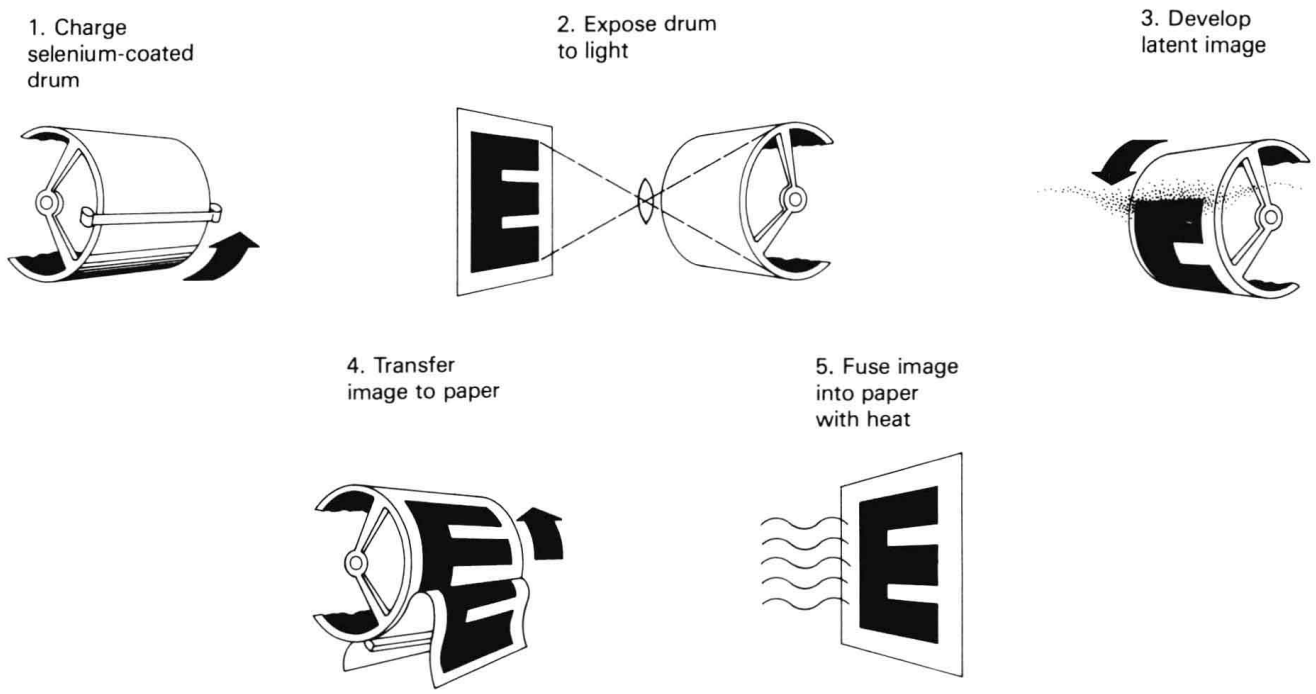
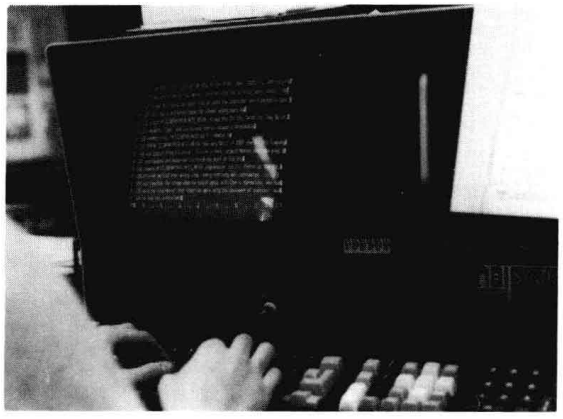


Figure 15.8 Diagram of Copier Operations Most dry-process copiers use the sequence of operations shown here. Larger machines use a variable enlargement/reduction lens between image and sensitized drum. (Photo courtesy of Xerox Corporation)

type on a sheet of paper at top speed without worrying about errors. The copy was simultaneously recorded on magnetic tape in standard **alphanumeric** (letters and numbers) code. Errors could be corrected by back-spacing and retyping. When the work assignment was completed, the sheet of paper in the machine was worthless as copy, but the magnetic tape contained a flawless magnetic facsimile of the original, edited and uniformly spaced. A magnetic tape reader (a transducer) then turned the magnetic impulses into a DC current pulsating in accordance with the magnetic pattern on the tape. This current could be used to control the mechanical operation of the same or another electric typewriter spewing out perfect copy at the rate of 150 error-free words per minute.

Since 1964 IBM and other business machine companies have steadily improved their technical equipment and expanded its usefulness in processing and storing data. Most dramatic has been the move away from a semimechanical, semielectronic approach to an almost entirely electronic approach. For example, the IBM Office System 6 features a cathode-ray tube display, high-density storage on a magnetic disc, a magnetic card reader and recorder, an all-electronic keyboard (keys are actually switches) that can handle 96 characters in a number of modern languages, and a high-speed ink jet printer capable of printing up to 92 characters per second. The whole assembly is easily **interfaced** (interconnected) with an easy-to-control computer that integrates data and word functions.



A secretary typesets a textbook page on an NBI System II business machine. Errors are easily corrected by electronic erasing and retyping of words or entire lines.

Robots and Robotics

The design and production of robots is the most rapidly growing sector of the Electronics Industry. The word *robot* refers to a machine that can to some degree move, work, and, to a lesser extent, "think" like a human. Its motion is limited to gliding along a bar or rail. However, a robot can move its flexible arm precisely and perform tasks that people consider drudgery. For example, the larger one-armed robots can do a rapid and thorough job of first priming the body of an automobile and then adding color coats of enamel. Teamed with other robots along an assembly line, a robot can quickly and precisely weld chassis and bodies together as they move at a constant speed by fixed stations. Smaller, usually two-armed, robots can rapidly assemble speedometers or carburetors from a supply of small parts or modules.

To perform these operations, a robot must have some degree of electronic vision (an electric eye) so that it can track its own movements and make corrections. A robot must also have an on-board microprocessor to control the hydraulic systems or servomotors that move the arm as the microprocessor is instructed by a master control panel built around a large computer that uses a magnetic disc drive for its software instructions.

Robotics, an all-inclusive term, refers to the design, manufacture, and effective application of robots. Japanese industries currently employ at least five times as many robots as American industries. The japanese have about 36,000 robots, largely in automotive plants, compared to about 6500 in the United States.

For more widespread application in industry and business, the present generation of robots needs improvements in the following areas: vision, tactile sense (sense of touch), and mechanical movement. The General Electric Company, one of the principal distributors of industrial robots, has recently developed a low-cost system of optoelectronics that permits visual information (at infrared frequencies) to be transmitted readily over an optical fiber of quartz, glass, or plastic to a light-sensitive diode, or detector (usually a charge-coupled device or CCD). After suitable processing, the diode's analog output voltage can be used to control precisely the movements of the robot's arm and any "fingers," or grippers, that are employed. GE has already developed a touch

As Chrysler K-cars move down the assembly line, robot arms make nearly 3000 welds on each body. Robots shown here are computer controlled and programmed from a nearby control center. Accordion-pleated sleeves on arms protect them from welding dust and metal particles. (Photo courtesy of Chrysler Corporation)

sensor system that enables robots used in assembling small parts into larger modules to reject nonstandard components.

Optoelectronics

Optoelectronics is the technology that deals with the coupling of functioning electronic modules by means of a beam of light. In the following discussion we focus primarily on the coupling of electronic units by the use of infrared light that travels over one or more optical fibers. Light travels along a suitable optical fiber as a DC electrical current travels along a metallic wire, but with several important differences: Light's passage does not produce a field around the optical fiber; the fiber is immune from interference from stray magnetic fields or noise fields through which it must pass; and the fiber can be used to transmit all types of data in digital or analog form. The most effective optical fiber is quartz, but it is also the most expensive. Highly purified glasses are somewhat less effective in light transmission but are much less expensive. Plastic fibers are the cheapest and are satisfactory for short runs. Propagation speed of light through an optical fiber is less than through free space.

Figure 15.9 shows the basic building blocks of a typical optoelectronics system. At the left, a data-bearing electrical signal enters the system, and the information it carries can be either analog or digital. The signal is applied across the light-emitting diode, D_1, and the diode emits a varying amount of infrared light varying at a signal frequency rate. In most low-cost systems the diode and a plastic lens are molded into the same plastic housing. The lens greatly improves the transfer of light from the diode to the end of the optical fiber.

Figure 15.10 shows General Electric's new series of optoelectronic components for low power. Both emitter and detector diodes are shown in their molded assemblies, which have special provision for close-coupling their lenses to the fitting that terminates the length of optical fiber used. A single fiber can be used or, more likely, the circuit will employ a carefully arranged bundle of small, individual fibers.

Once the varying light beam has been started down the optical cable, it will travel, at appreciably

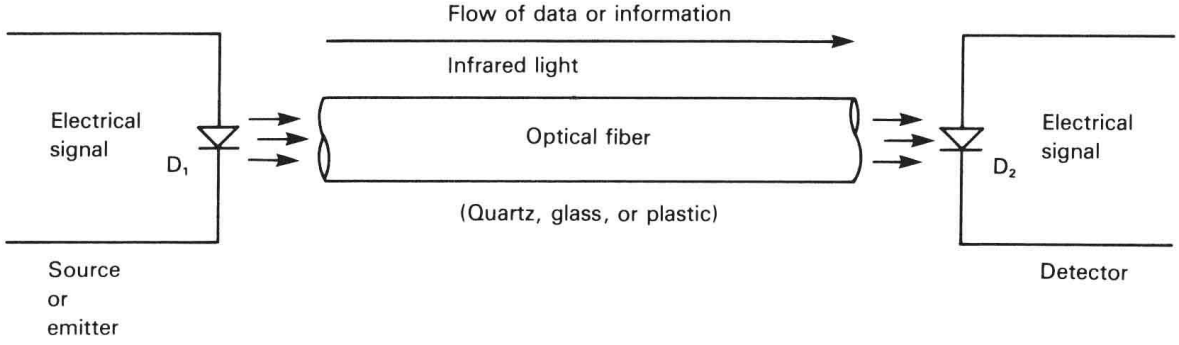


Figure 15.9 Basic Optoelectronics Circuit When an analog or digital voltage is applied to diode D_1, it emits infrared light that varies with the DC signal. A lens concentrates this light on the end of the optical fiber through which the light travels at less than 186,300 mps. At the end of the fiber, the light is focused by a lens on the detector diode D_2. The output of this diode is an analog or digital signal having the same waveform as the one originally applied to D_1.

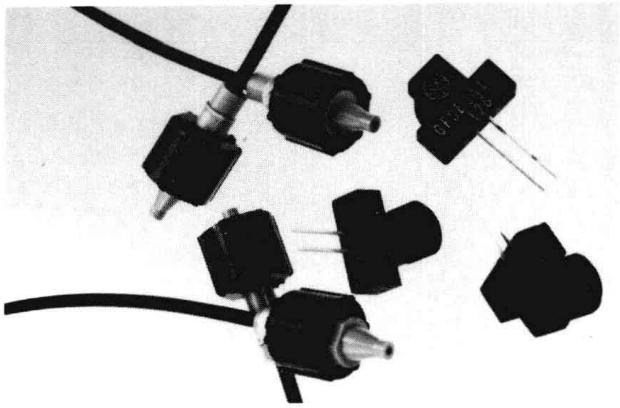


Figure 15.10 Optoelectronics Components An assortment of low-cost fiber-optic devices are available from General Electric for use in short-haul data transmission. Shown are three semiconductor light sources and detectors designed to mate with plastic ferrule connectors that readily couple to over forty types of optical fibers. (Photo courtesy of General Electric)

less than the free-space speed of light (186,300 mps). Arriving at the far end of the cable, it will be focused by the plastic lens on the sensitive detector diode that acting as a transducer, transforms the infrared light variations back to either analog or digital electrical impulses.

Optoelectronics systems employing components such as those shown in Figure 15.10 and operating at the very modest power level of 100 microwatts can achieve reliable, wide-bandwidth transmission of data up to a distance of one kilometer, or about 3280 feet. The efficiency of such a system and its bandwidth is greater than a system using conventional, twisted-pair conductors, coaxial cable, or microwave guides. Now that General Electric has brought out a line of low-cost emitter and detector diodes, along with plastic fittings that permit rapid attachment of optical fibers to plugs that attach to the emitter and detector assemblies, design engineers will be inclined to use this optoelectronics system for data transmission when distances are relatively short. For example, the system can be used in linking computers to peripheral equipment and remote terminals, in data transmission networks, in communication systems in military aircraft, and in the interconnection of sensors, microprocessors, and supervisory microcomputers in new automobiles. Telephone companies use a powerful laser beam, along with booster amplifiers, to extend the distances of communication.

The only serious disadvantage to the use of an optoelectronics system is that it requires the use of two transducers—one as an emitter, the other as a detector. Since the efficiency of any transducer is always less than 100% (usually much less), overall system efficiency can never be very high. However, the system's effectiveness for wide-band data transmission over relatively short distances and its immunity to interference from electromagnetic and noise fields are compelling advantages.

LATEST TRENDS IN ELECTRONICS

By 1984 certain trends were evident in the home entertainment field. Video discs and video disc players were selling briskly as prices dropped, and viewers were quick to respond to the appeal of prepackaged programming with color and stereo sound. Only VCRs, with an improved stereo sound channel (a product of processing the aural information along with the visual), were selling more rapidly. The video game industry was experimenting with a fusion of conventional digital games and the superb visual and audio quality of video discs. (Consumers have rebelled against the matchstick figures and poor sound of the present generation of video games.)

In the personal computer industry the trend is toward buying the best computer you can afford and buying from an established manufacturer with the largest possible spectrum of software. The quality of peripherals is also given high priority.

The home market has seen a proliferation of wired and wireless telephones and a corresponding increase in automatic telephone answering and dialing devices. Increased sales of telephone equipment are due in large part to the breakup of the Bell System, which had opposed equipment interconnection.

In business offices the greatest use of electronic equipment focuses on the handling of data. Office-size computers are frequently interconnected to other computers. Computer terminals are linked to a very large main-frame unit, housed elsewhere, so that individuals at their desks can communicate with each other or can interrogate the main computer. Multicolor copiers and duplicators spew out flawless copies at high speeds, and word processors produce professional-looking brochures and manuals.

In industry many of the electronic modules employing active devices have been joined by microprocessors—miniaturized computers mounted on a single silicon chip. Like a real computer, a microprocessor has a memory to provide it with just enough artificial intelligence for the job at hand. Its output is usually analog, or conventional current or voltage suitable for operating solenoids, synchromotors, and other industrial control devices. When more sophistication is required, a second microprocessor may be used to supervise the operation of the first. Finally, in the fields of robotics and optoelectronics, tremendous expansion will have occurred by 1990.

SUMMARY

The dominance of the radio console continued until the late 1940s, when it was challenged first by FM broadcasting and then by TV broadcasting. By the late 1960s, transistors were rapidly replacing vacuum tubes. By the late 1970s, ICs were replacing transistors. Stereo amplifiers of wide frequency response and large power output became available at low cost. Color TV sets, already of remarkable fidelity, increased their impact on viewers through the use of screen projection technology. Microprocessors, with their chip-mounted artificial intelligence and control functions, found new applications in home laundry units and microwave ovens. Cordless telephones and message recorders proliferated with the breakup of the Bell System. Video games burgeoned and then lost public support. Personal computers followed a similar curve. VCRs, offering time shifting of favorite TV programs, became best-sellers, closely followed by video discs, which offer quality programming on a prerecorded basis.

In modern offices the most important applications of electronics relate to the collection, processing, and management of information and data. Computers and their terminals play the dominant role, and their use will be expedited by the newly developed technology of optoelectronics, which uses data to modulate a light beam transmitted over optical fibers of quartz, glass, or plastic. Data can be transmitted for thousands of feet with little chance of interference from electrical equipment or stray magnetic fields. High-speed electrostatic copiers have also reached a high level of performance and reliability, and many of the larger machines use a form of readout that alerts the repairer to the cause of a malfunction. Word processing, in which data in alphanumeric form are digitally encoded, processed, and stored on magnetic cards or discs, has reached a high degree of sophistication. After processing, perfect copy can be recovered in printed form at high speed.

The widespread use of electronics in manufacturing has been made possible by a variety of microprocessors, supersmall computers on chips. Microprocessors respond to data supplied in digital form. They have some memory and can be readily teamed with other microprocessors or even with computers to control a specific function of a moving arm or lever. Used in this way, they control industrial robots. Robotics promises to be a rapidly growing field as robots acquire better vision, more sensitive touch, and greater mechanical mobility and flexibility.

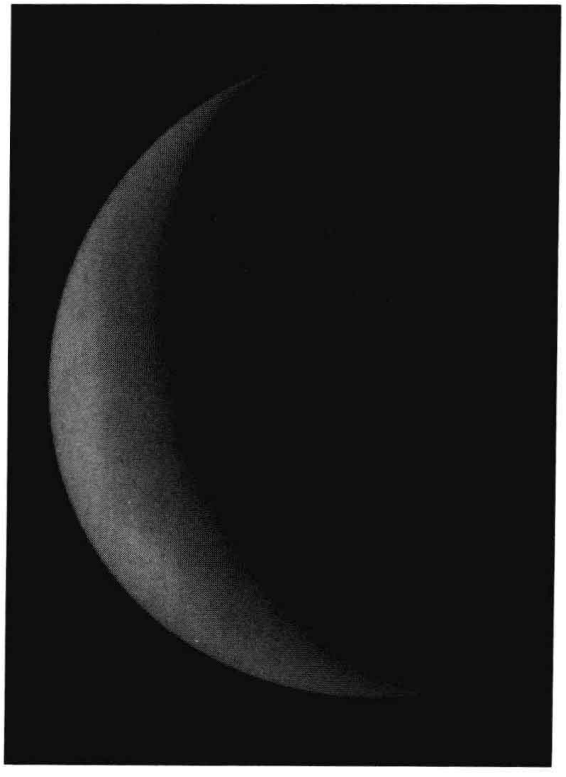
KEY WORDS AND CONCEPTS

acrylic lens
alphanumeric
anemometer
electrostatic copying
Fresnel lens
interfaced
interlocks
ionization chamber
magnetron
microprocessor
optoelectronics
photoelectric cell
pixels
robotics
stirrer
strobe flash tube
varistor
waveguide
word processing
xerography

SELF-TEST QUESTIONS

1. Explain how microwave energy can be used to cook food.
2. How can a microprocessor linked to a microwave oven simplify food preparation?
3. Why have synthetic fibers and permanent-press clothing stimulated the sales of washer-dryer combinations with microprocessor controls?
4. Describe the operation of an ionization chamber smoke detector.
5. In inexpensive automatic cameras, is it the shutter speed or the aperture opening that is varied by the electronic circuitry?
6. What are two advantages in the use of silicon photocells for camera control?
7. How does Polaroid use ultrasound in the automatic focusing of its new Sonar camera?
8. If an anemometer is actually a simple AC generator spun by the wind, the indicating instrument must be a specially calibrated form of what common electrical meter?
9. Explain how General Electric can use a standard color TV chassis with a 13 inch picture tube to fill a viewing screen measuring over 1000 square inches.
10. How do telephone answering systems use the telephone's ringing signal to place the playback and recording unit in operation?
11. On what law is the operation of an electrostatic copying machine based?
12. IBM's Selectric typewriter records ordinary copy on magnetic tape in the form of alphanumeric code. How is the coded information turned back into typewritten copy?
13. Newer word processing equipment usually makes use of a TV screen that provides a constant record, including revisions, of what is being recorded on magnetic cards or discs. Why is this visual display of great assistance to the operator?
14. What is meant by robotics?
15. Describe an industrial robot.
16. What is meant by optoelectronics?
17. Why can an optoelectronics data-transmission system never be 100% efficient?

16 A Look at the Future of Electronics



This photo of the cloud-shrouded planet Venus was taken by an approaching NASA space probe. (Photo courtesy of NASA)

Five-Year Forecast for Electronics
- Industrial Production
- Industrial Employment

Five-Year Forecast for New Applications
- Medicine and the Health Occupations
- Space and World Exploration
- Telecommunications
- Audio Systems
- Video Systems
- Electronic Musical Instruments
- Computers
- Calculators
- Electronic Clocks and Watches
- Power Generation and Distribution
- New Uses for Electronics in the Home
- New Uses for Electronics in Business and Industry
- Automotive Electronics

In the preparation of this chapter, the most recent Department of Commerce predictions were studied, trade magazines were scrutinized and notes taken, and people with close ties to West Coast electronics industries were interviewed. Making five-year forecasts involves the risk of being in partial or complete error, but it is a risk that must be taken.

This chapter is in two sections. In the first we look at the future of the Electronics Industry as a whole. In the second we look at specific areas and applications. In format this chapter follows the topical sequence used in the applications chapters (Chapters 8–15).

FIVE-YEAR FORECAST FOR ELECTRONICS

The forecast that follows focuses on both industrial production and industrial employment because the former makes the latter possible. Industrial production eventually has an impact on employment in the equipment-servicing trades and in specialized jobs like field engineers and factory representatives.

Industrial Production

The Department of Commerce forecast for 1982–1983 predicted total sales of Electronic Equipment (both Government and Consumer Electronics) would reach $18.5 billion, and growth over the previous year would be near 10.1%. This growth was taking place during an economic downturn, and the primary reason for the good showing was that the government sector of the industry was fueled by large Department of Defense contracts. At this writing, it is too soon to be positive that the DOC forecast was completely accurate.

The DOC forecast for the same period for Electronic Components—$23.9 billion in sales and a growth rate of 8%—seems very close to the mark. However, it may not have reflected a late-in-the-year weakness in IC sales caused by overproduction and increased imports.

Until 1980 the DOC made separate forecasts for production and employment in each of three fields: Electronic Systems and Equipment (large items principally sold to government agencies); Consumer Electronics (electronic equipment items principally sold for home use); and Electronic Components (items sold both to the firms that assemble the two categories of electronic equipment and to technicians who repair and service this equipment). Since 1980, the Department of Commerce has grouped Electronic Systems and Equipment and Consumer Electronics in a single, all-inclusive category titled Electronic Equipment. Trade magazines, largely because of their consumer orientation, still prefer the original three categories.

Electronic Equipment will continue its steady growth in production and sales, and this growth will probably peak during the 1985–1986 fiscal year, when the impact of defense spending will be felt to the fullest degree. Congress has already appropriated money for many defense-related projects, and a great deal more is probably forthcoming. Furthermore, because of the classified nature of much of the electronic equipment manufactured for defense, all production takes place in the United States. There is no danger that the manufacturing operations might be exported to another country to save money.

Consumer Electronics will be making a somewhat smaller contribution to this growth, primarily stimulated by items like VCRs, video discs and players, and possibly the more expensive personal computers and their peripherals. Evidence indicates that the typical consumer not only has money to spend but again is willing to spend it, even for big-ticket items.

Industrial Employment

Electronic Equipment Production

Employment in the area of electronic equipment production for government and large businesses is sure to show a steady increase because of the vast sums already appropriated for defense. NASA also has large sums earmarked for defense-oriented items

such as surveillance and communications satellites. Production workers engaged in this area of large electronics systems manufacturing will generally enjoy higher wages and more job security than workers engaged in Consumer Electronics.

Consumer Electronics

Consumer Electronics will show a slow but steady growth in employment, principally brought about by the economic resurgence following the 1981–1982 depression and the willingness of consumers to spend freely once more, even for big-ticket items such as computers and video disc players. However, the outlook is clouded by unpredictable moves on the part of large employers—for example, the export of manufacturing to the Far East to meet the price competition of foreign products arriving in the American marketplace. At present this practice is limited, but almost all manufacturers of Consumer Electronics engage in it to a certain extent to reduce the final cost to consumers.

Early in 1983 sales of video games and video game players first slowed and then dropped sharply as the novelty wore off and consumer resistance stiffened. Companies like Mattel and Atari found themselves near bankruptcy and closed down their U.S. production lines. Atari exported its game manufacturing to the Far East to take advantage of lower labor costs. A somewhat similar fate awaited personal computers (PCs). By mid-June 1983 manufacturers were shipping 450,000 new computers each month to dealers who were unable to sell more than 275,000 units per month. In the resultant crisis Texas Instruments closed their PC division, and many smaller companies went bankrupt.

Early in 1984, stability seemed to have returned to Consumer Electronics. Large-screen color TV sets were selling well, although the list price was too high for mass sales. Manufacturers such as RCA, who have opted for a more brilliant image by the use of a high-power projection system and a separate screen, were also experiencing good sales. The forecast is for prices of these items to dip and sales to surge ahead.

Video cassette recorders (VCRs) are moving briskly as prices have dropped sharply and their market has been greatly expanded. Newer models feature stereo audio systems, which have further increased sales. Prices will probably stabilize, but sales will continue to improve.

The digital audio disc (DAD), or the compact disc (CD), represents the first breakthrough in audio recording in more than thirty years. However, the digitally recorded and reproduced disc is still well in the future because new and very expensive recording equipment must be installed both in the recording studio and in the home. Sales of these discs will therefore be limited over the next few years.

Among the newest developments, Eastman Kodak Company plans to enter the home video market with a new camera that records on 8 millimeter videotape. (Kodak has for half a century pioneered in the development of a wide spectrum of cameras designed for use with silver halide films, which Kodak developed.) IBM has also entered the personal computer field with a rather expensive home computer and the necessary peripherals as well as a wide spec-

Sony's compact disc player uses a 4 3/4 inch digitally recorded audio disc on its turntable. The laser-recorded, grooveless disc is read by a low-powered laser. The digital signal recovered is processed into an analog stereo signal ready to be plugged into a home audio system. (Photo courtesy of Sony Corporation)

trum of compatible software. Manufacturers of other PCs are considering conforming to IBM's software standards—a move that would simplify the problem of lack of compatibility between the hardware of one manufacturer and the software of another.

Electronic Components

Slow, steady growth in employment will occur in the field of Electronic Components, which supplies components for the manufacturing of government- and consumer-oriented equipment and repair parts for equipment servicing. The driving force will again be the funds already appropriated for the construction of defense-oriented electronic systems. One factor contributes to uncertainty in this field: Until very recently the japanese had not achieved large-scale production of the more complex ICs, and the market belonged almost entirely to U.S. manufacturers. Today the market is flooded with many Japanese ICs at prices so low that the ability of U.S. manufacturers to compete is in question.

Electronic System Servicing

The number of technicians employed in servicing and maintaining electronic systems for government agencies will remain relatively stable or will decrease slightly because of the freeze on civil service hiring and, to a lesser degree, because equipment channeled to the armed forces is serviced by technicians who have graduated from the military's own training schools.

The number of technicians engaged in servicing consumer products is expected to show a slight decrease for two reasons. First, the level of expertise required to service sophisticated electronic equipment is so high that many people who have worked for years in the field are unable to handle the advanced circuitry and the complex test equipment required. Second, the cost of repairing a piece of consumer-oriented electronic equipment, such as a TV receiver, may be reaching the point where purchasing new equipment is cheaper than extensive repairs. In the area of pocket radios, pocket tape players, and simple calculators, we have already reached the point where we throw them away rather than repair them.

FIVE-YEAR FORECAST FOR NEW APPLICATIONS

Topics discussed in this section are in the same sequence as they appear in the preceding applications chapters. If there is one thread running through all the discussions that follow, it is the ever-increasing use of digital technology in electronic areas that previously made no or little use of it. In every case the result is increased flexibility and effectiveness, although almost invariably at the price of increased complexity and expense.

Medicine and the Health Occupations

In general, computers will be used extensively to collect and organize information in understandable form (as is already being done by machines like the CAT scanner). There will be increased emphasis on the use of ultrasonics for partial and full body scans since this technology is effective and poses little or no hazard. When X-rays and other forms of radiation are used, microprocessors will keep the degree and length of radiation within exact limits to eliminate the possibility of overradiation.

Thermography will show some increased use because of its effectiveness in locating and diagnosing circulatory problems near the body's surface. Thermography uses a system that is basically a closed-circuit TV camera and monitor responsive only to infrared (heat) radiation. The newer machines show body temperature changes in color and make the pattern on the monitor screen easier to interpret.

The newest development in diagnostic imagery is a technique called NMR, or nuclear magnetic resonance. It differs from the CAT scan in that it uses

noninvasive magnetic fields rather than tissue-damaging X-rays. For the first time, clear and detailed pictures of the human brain and surrounding tissue can be taken. Eventually NMR scanners will replace CAT scanners but will cost well over $1 million for each unit.

In therapy, ultrasound is finding wider use now that more effective transducer units have been devised. Ultrasound can produce the same deep therapy previously obtained by the now-obsolete diathermy units, which posed the hazard of excessive electromagnetic radiation. Such treatment is often referred to as micromassage.

Space and World Exploration

NASA had a remarkably successful year in 1983. First, its budget was increased. Second, space probe Pioneer 10 early in June crossed the orbit of Neptune and permanently left the solar system—the first man-made object to do so. (Being nuclear powered, it will communicate with Earth from interstellar space.) Later in the same month space shuttle Challenger marked up its seventh successful shuttle mission by launching two communications satellites while in orbit. Voyager II is scheduled to photograph Uranus in 1986 and Neptune in 1989. Meanwhile, NOAA has further increased its network of FM stations broadcasting local weather on about 162 megahertz and has placed several improved geosynchronous and orbiting weather satellites into orbit. In his January 1984 State of the Union address, President Reagan announced that he had asked NASA officials to start planning for a permanent space station to be operational within this decade. Contracts for research and development of sophisticated electronic systems with private organizations by both NASA and NOAA will probably continue, but at somewhat lower levels of activity.

Telecommunications

The most exciting current development in the field of broadcasting and telecasting is the distinct prospect of improved audio quality for both media. In the case of conventional AM broadcasting, the FCC has under consideration several approaches to stereo broadcasting that would provide markedly better sound for those who purchase new stereo equipment. People still using older receivers would hear the same program in monophonic sound and with no loss of existing quality.

Improvement in TV audio will come in two steps. The first step was taken by AT&T recently when it substituted audio network lines for lines previously in service. The new lines are capable of high-frequency response to 15 kilohertz, compared to an old high of about 5 kilohertz. The telecasting industry is urging the FCC to study the possibility of stereo FM for television sound. Manufacturers seem generally in favor of this move, which would involve much more sophisticated audio amplifiers and loudspeakers in new TV receivers. The increased cost would need to be reflected in a higher purchase price. In March 1984 the FCC started processing AM and TV station requests for permission to use stereo transmission.

This amateur radio operator, using the latest in single-sideband equipment designed for use in the amateur bands, can contact stations throughout the world. (Photo courtesy of the Heath Company)

Radio amateurs may have to share some of their allocations with other radio services but may gain additional bands between their existing assignments if proposals scheduled to be made in an upcoming world conference are approved. It seems likely that citizens band operators will be moved from their present 27 megahertz allocation to a new and more commodious one at about 900 megahertz. There would be no interfering skip at these frequencies, and highly effective antennas, just a fraction of the size of those currently in use, would make mobile operation much simpler. However, all existing CB transceivers (perhaps 30 million, both licensed and unlicensed) would become obsolete overnight. In 1983 the newly approved 10 megahertz amateur band was in use. Code transmissions only were permitted.

AM stereo, operating under special FCC authorization, is now being heard over a few of the AM stations in the larger cities. Unfortunately, very few AM listeners were provided with the special receiving equipment necessary for the stereo effect.

TV stereo is still in the future. Cable TV is making serious inroads on the audiences of network TV stations. Cable-ready TV sets with built-in cable converters are available at any appliance store. All the commercial radio and TV networks, as well as PBS, are connected by satellites and ground stations in place of microwave links, and the improvement in both aural and visual quality is remarkable. Within the next five years the audience of standard TV stations and networks will become increasingly fragmented, and this development will have a marked economic effect on TV broadcasting.

Audio Systems

In the recording studio, digital technology will be used increasingly each successive year in preparing the master tape and the master disc. The result will be improved signal-to-noise ratio and a much greater volume range on the processed tape or record. This technology will be too expensive for the audio hobbyist, but the improvement will be seen when a prerecorded tape or disc is played back.

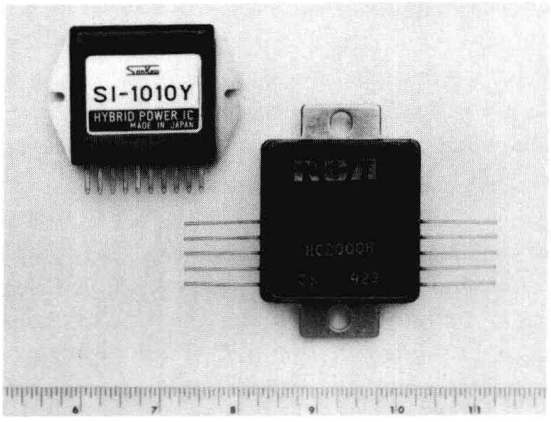
Metal-particle tape coatings will so improve the frequency response of cassette-type audio equipment that it will render reel-to-reel machines obsolete except for professional work in radio stations and recording studios. Cassette quality will equal, or possibly exceed, that of reel-to-reel equipment of just a few years ago.

The weakest link in the home audio system has always been between the loudspeakers in their enclosures and the air of the room. Room acoustics plays an important role in the sound of a home system, and each installation has unique problems due to room size and to unpredictable amounts of reflections and absorption. Many audio enthusiasts have as their goal the realization of concert-hall presence in the living room, but cranking up the volume control or adding multiple speakers seldom helps. Synthetic reverberation, or time delay, may help create the desired effect. Currently, integrated circuits to provide such time delay are available, and future audio systems will certainly make provision for this feature. The result will be some degree of control over living-room acoustics and a reasonable simulation of presence.

During the next five years more of the discrete components used in assembling audio systems will disappear into large-scale integrated circuits (chips) that will offer both advantages and disadvantages. On the plus side, the amplifier will be smaller and more compact and will require a lesser amount of expensive labor to assemble. On the negative side, if even one tiny component of the LSIC fails, the whole complex and expensive unit must be replaced at a cost of $40 to $50. Also, FET transistors will replace the bipolar types.

As engineering costs are gradually written off, the cost of both digitally recorded audio discs and their special playback units will fall to the point where a mass market will develop. All premium singles and albums will be both recorded and played back through the use of digital technology.

Eventually the pinpoint laser beam will probably scan the surface of all phonograph records and replace the diamond stylus, which wears away the



Each of these modules (VLSICs) is the complete push-pull amplifier unit for one L or R channel of a home stereo system. Each contains a voltage amplifier, a power driver, and a heavy-duty power-amplifier stage. The Sanken unit is rated at 30 watts, the RCA at 40 to 50 watts. As is typical of IC-based electronic equipment, if one component fails, the whole module must be replaced.

This couple is enjoying the exceptional stereo audio quality as well as the visual programming on a video disc played back on a late-model machine. (Photo courtesy of Pioneer Video, Inc.)

delicate grooves embossed in the vinyl disc. We already have the technology to use laser beams; the question is whether listeners are willing to absorb the additional cost. In theory, records played back in this manner would never wear out or show scratches and pops caused by the mechanical stylus.

Video Systems

We have already discussed the future of VCRs that use magnetic tape cassettes and video disc playback equipment that uses a prerecorded vinyl disc. These two items will represent a substantial part of Consumer Electronics sales within the next five years. However, the problem of compatibility must be overcome before this goal can be realized. The problem is even more serious in the video disc field, where RCA is using a mechanical stylus, for example, while Magnavox's more expensive model uses laser beam scanning and a different turntable speed.

All top-of-the-line VCRs now have full stereo capability and greatly improved audio. A few innovative manufacturers now use the input video and audio signals from the TV tuner to frequency-modulate an RF carrier wave prior to the tape recording process. During playback this wave is recovered and demodulated. The result is such high visual and aural quality that the signals are indistinguishable from the original.

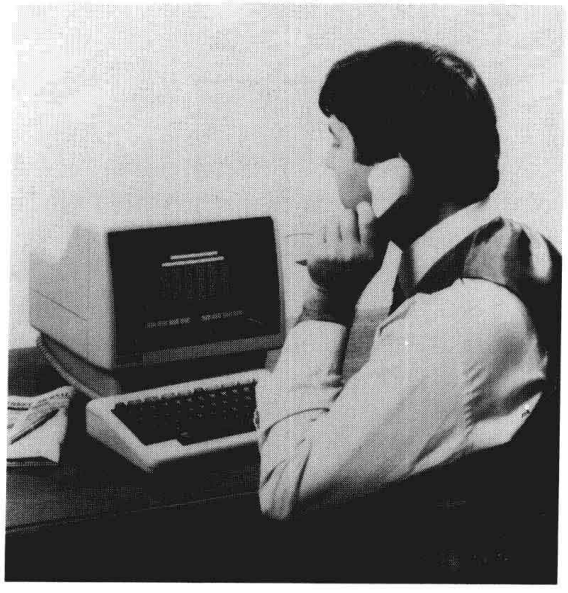
In the video disc area two basic types of playback equipment will be sold: a deluxe line that will use laser beam scanning of the record grooves and a lower-priced series that will use mechanical scanning. Eventually all manufacturers' discs in each category will be compatible.

Electronic Musical Instruments

No great changes are forecast in electronic musical instruments for the next five years. ICs will make possible more effective controlled reverberation with

consequent improvement in tonal quality. The same ICs, when used as wave-shaping devices (formants), are capable of producing all the traditional organ tonal families plus hundreds of new ones. Some experimentation is under way with plug-in formant modules. Such an approach could provide a relatively simple and inexpensive organ with many of the tonal facilities offered by the much more expensive concert models.

The newer models of electronic organs are easier for the nonprofessional to play. Some changes are purely mechanical, such as rearrangement of the various stops so that all similar tonal families are grouped together or inclusion of preset buttons that bring a number of stops into play at the flick of a thumb. The percussion stops (piano, orchestra bells, and drums) are much more realistic and will probably show steady improvement in the next five years.



Hewlett-Packard's new 9 inch office display terminal occupies less than a cubic foot of desk space. It can be used in office operations that require data entry, data inquiry, and development of new programs. (Photo courtesy of Hewlett-Packard)

Computers

Many microcomputers, or personal computers, have been sold for home use since 1980, but the number of annual sales is expected to slowly decrease during the next five years as people realize the importance of a large investment in peripheral equipment items for tasks such as word processing. Peripherals, or terminal equipment, include as a minimum floppy disc memories and cathode-ray tube displays.

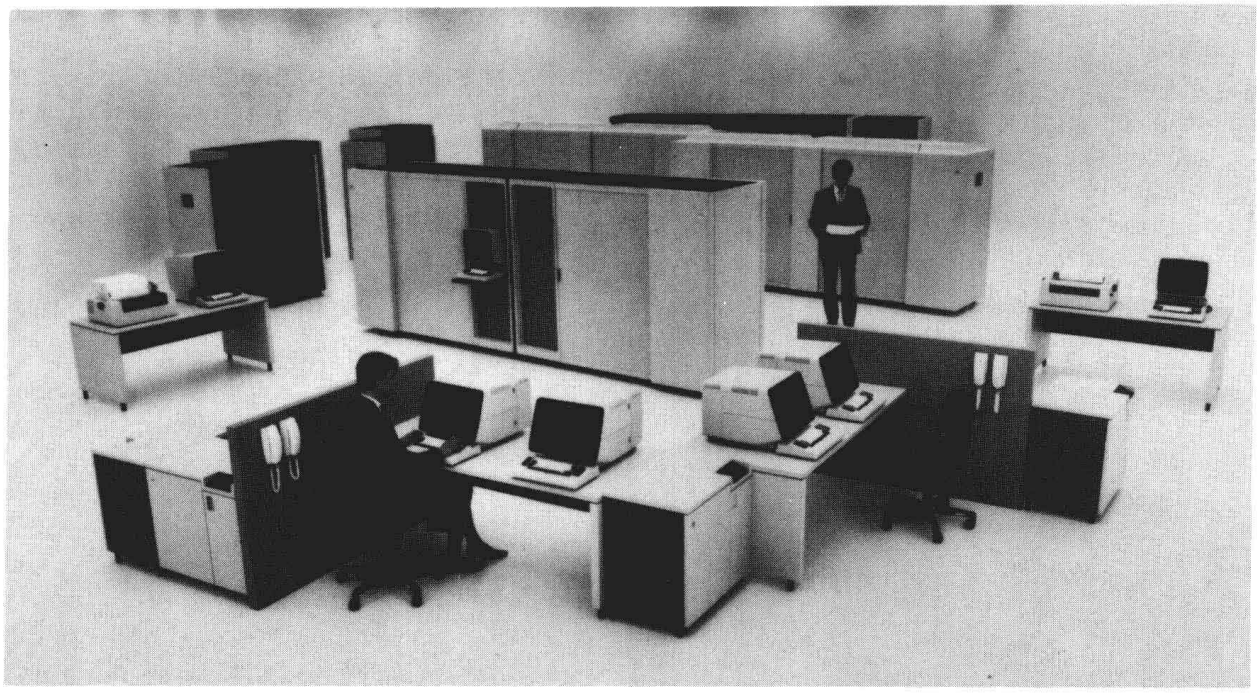
Predictions that home-based computers will soon balance checkbooks, figure income tax, prepare meals, meter out water for lawns, and generally run an entire household with machinelike precision are probably unrealistic. All these operations are possible, but many are completely impractical because of their cost and complexity. The minimum acceptable memory for personal computers is now 64K bytes.

In the commercial computer field, look for a continuing reduction in physical size and complexity, largely due to more extensive use of ICs. Look, too, for a prodigious increase in the memory capabilities of commercial computers, which will use highly specialized ICs for data storage. This memory will be permanent until erased—not volatile, or lost, when the power is turned off.

Although it may seem remarkable that a large computer can make a multidigit computation in a microsecond or less, engineers are always striving to decrease this time even further. If light and photon movement could take the place of the comparatively slow transit time of electrons through transistors, the computation time could theoretically be reduced to one-trillionth second. This achievement would require development of an active device that would operate on light waves and photons rather than on the movement of electric charges. Reportedly, such an analog has been constructed. If the prototype can be reproduced, it will represent a remarkable breakthrough.

Calculators

Digital displays will eventually use far less current than today's LEDs and LCDs and thus make possible



This IBM model 3084 is the firm's most powerful computer. Facilities include 4 CPUs, 64 million characters of main memory, and 48 channels. (Photo courtesy of International Business Machines Corporation)

both size reduction and increased battery life. In 1983, LEDs had been almost completely replaced by LCDs. Batteries now last up to a year and a half. More memory will be included so that calculators will be able to store figures and data for later recall and use.

Electronic Clocks and Watches

Improved digital displays in electronic clocks and watches will be easier to read under adverse lighting conditions, and their decreased power consumption will further extend battery life beyond the present expectancy of about one year. Battery life is particularly important in watches, where batteries are essential as a power source. Petite electronic watches, almost completely absent from today's marketplace, will soon be available as new time-keeping modules are designed and smaller batteries are manufactured.

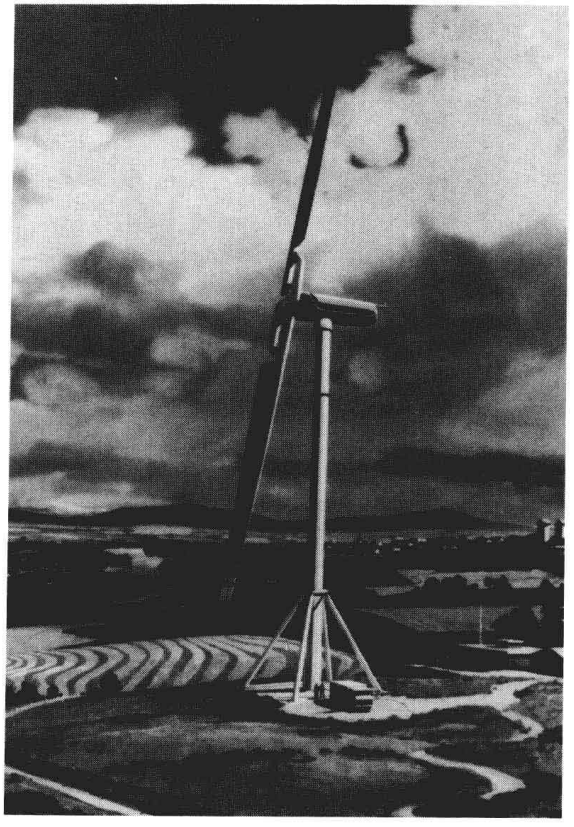
Digital watches are often difficult to read at a glance. By late 1985, digital watches will have hour and second hands made up of a liquid crystal digital display that moves across the watch dial. Because of their increased complexity, such watches will be more expensive than those with a simple, one-line digital display, but they will be much easier to read at a glance or under adverse lighting conditions.

Power Generation and Distribution

Electrical power is one of the oldest applications of electricity and electronics, and no significant developments are expected in this area within the next five years. Environmentalists have been trying to block new starts of conventional (fission) nuclear plants and to halt the operation of the nation's largest

nuclear plant, the Trojan generating facility in Oregon. Scientists have already demonstrated atomic fusion using tremendously powerful magnetic fields to provide the pressures required and laser beams to supply intense heat. However, atomic fusion on the earth's surface as a source of heat to produce electrical energy is at least as far away as the year 2000.

Scientists have developed a new solar cell with a conversion efficiency of solar-to-electrical energy of about 30%, a distinct improvement over regular silicon cells. (Stanford Research Institute scientists announced late in 1979 development of a cheaper method for producing large amounts of silicon for solar cell manufacture.) However, electrical energy obtained directly from solar energy is very expensive and will play no significant role in the economy for many years. Similarly, wind power as an electrical energy source will be of little importance for at least the next decade.



In this drawing of a proposed wind-power generation site, a propeller-like rotor spins the shaft of an AC generator. (Photo courtesy of NASA)

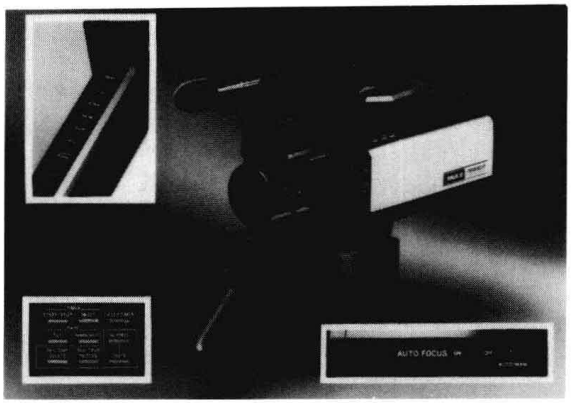
In 1983 the Department of Energy decided that the completely solar-heated home was possible but impractical at latitudes much farther north than 40° to 45°. Farther north the payback period is too long. However, passive systems of heat collection and water heating are economical even in the most northern tier of states.

DOE's joint project with the Bonneville Power Administration in the Columbia River gorge near Goldendale, Washington, continues to explore the comparative advantages of large, wind-driven generators rated in megawatts and smaller wind machines rated at a few kilowatts. So far, the small units intended to supplement the electrical power needs of single homes have proved efficient, economical, and reliable. The giant wind turbines, capable of serving a large number of homes, are out of service much of the time.

New Uses for Electronics in the Home

As previously noted, in the next five years increasing sales are expected in the home market for VCRs, video disc players, microwave ovens, microcomputers and their terminal equipment, and large-screen color TV sets of the projection type. Microprocessors, already at work in the home in controlling microwave ovens and washers and dryers, will soon be applied to home heating plants, electric stoves, and refrigerators in the interest of reducing the amount of energy consumed.

Cameras with sophisticated electronics will shortly use optical principles rather than ultrasound to provide instantaneous and exact lens focusing. A battery-powered DC motor will move the lens barrel to the precise setting. The principle used, in general, is that a properly focused image provides the greatest



RCA's newest video camera, the Newvicon, has many of the features of the latest generation of film cameras—for example, power-zoom lens and ultrasound self-focusing. It also has electronics for previewing material just recorded. It is designed for the advanced amateur. (Photo courtesy of RCA)

light intensity, and this light can be metered electronically. The technique is already at hand to take snapshots on magnetic tape and make immediate optical prints in full color. In four or five years this technique should become more widespread.

Newly available is an electronic language translator that resembles a commercial portable cassette player. The unit has a slot to insert a tape cassette and has perhaps two dozen plastic push-button switches on the front panel, each with a designated phrase in English. When a cassette of any given language is inserted and the button pushed for the desired phrase, a well-modulated synthetic voice speaks the correct words. Electronic translators will prove more valuable as their vocabularies are extended. They will be useful not only for international travel but also at all levels of education for applications such as vocabulary development. Language translators are expected to sell for around $500.

Personal computer sales will grow more slowly in the next five years, but the number and quality of computer peripherals will increase. Expect improved software to make possible greater use of home computers in learning. Also, expect owners to make better use of word-processing technology. Look for video games to make use of video disc technology to improve visual and audio quality dramatically.

New Uses for Electronics in Business and Industry

The most dramatic change in business and industry will be the widespread application of basic microprocessors to hundreds of thousands of machines used to manufacture, test, and package many industrial products. Microprocessors are closely related to microcomputers, but they differ in two important particulars. Although a microprocessor may use a silicon chip for its central processing unit, it has far less memory potential, and its input and output circuits are specifically designed for the control of other circuits rather than for complex data handling.

In manufacturing, at each test position a microprocessor will monitor, check, and supervise all aspects of production and assembly. For example, a unit might be installed in an automotive factory to determine whether all carburetors entering the assembly line are functioning at maximum efficiency. Eventually, microprocessors will be found at all major test positions and will have wide application in factories for testing and controlling industrial pro cesses. Microprocessor use will be widespread because the units are inexpensive, reliable, and easy to apply to the task at hand. Even more important, they will produce a uniformity of quality that has never been achieved before.

The larger aircraft already use scores of microprocessors and microcomputers to make their operation easier, more economical, and less prone to human error caused by confusion and fatigue. Practically all the instruments in aircraft control panels now use digital readouts, which involve the use of sensors, analog-to-digital chips, and microprocessors. The display usually employs LEDs of varying colors.

Automotive Electronics

Electronic ignition, widely employed in substantially all U.S. automobiles since the early 1970s, was soon followed by the first of the fuel injection systems. By the standards of the 1980s, the first systems were elementary and lacking in sophistication. In 1984 models the fuel injection system is microprocessor coordinated with other engine functions to provide optimum engine economy and performance. For example, in the 1984 Ford Tempo a sophisticated onboard microprocessor module (which Ford calls its EEC-IV computer) provides both supervision and coordination between the microprocessors assigned to specific engine functions, such as monitoring fuel-air mixture, spark advance, intake-air temperature, engine speed, engine coolant temperature, and the amount of O_2 in the engine exhaust.

Late in 1983 Toyota announced development of a suspension system that uses a microcomputer to adjust the damping forces to meet the rapidly changing demands encountered in high-speed driving. Both speed and steering sensors relay encoded data to the series of microprocessors that control suspension damping as speed and steering factors rapidly change. Mazda has similar but less sophisticated controls scheduled for some of their 1984 models.

In more expensive models car manufacturers have replaced the instrument panel lights that warn of low oil pressure and engine overheating with elaborate LED readouts that provide more information and are more attractive. Readouts require a sensor, an integrated circuit, and an LED display—all relatively low-cost items in today's electronics marketplace.

The tinny, synthesized voices of the early 1980s that advised you that the fuel or oil supply was low or the engine was overheating have been replaced by a highly specialized IC developed by Texas Instruments. Auto dealers can supply the proper IC for either a male or a female voice. The new ICs even incorporate expression so that the voice sounds animated and lifelike.

The Texas Instruments chip that now provides synthesized voice will play an important role in development of road warning systems. Outboard sensors mounted on the car's body will relay encoded digital information from very low frequency, low-power radio transmitters strategically located along the highway. When the encoded data reaches the chip, a voice from a special loudspeaker will advise the driver of hazards ahead.

The technologies involved in road warning systems are already known to electronics engineers, and the equipment is available. However, installation of such a system on the nation's major highways would be so expensive that it could not be completed, even on the interstate system, until 1990 at the earliest.

In conclusion, if any one theme best describes the overall trend in electronics in the decade ahead, it is certainly the expansion of digital technology into new areas that will be dramatically improved by its application. Digital technology has already revolutionized audio recording and space communications and shows great promise in video recording and in television. The teaming of computers and microprocessors with a variety of electronic equipment will characterize the electronics of the future.

Index

Abacus, **239**
AC, **71**–81
  applications of, 71
  DC conversion to, 265–66
  DC versus, 72–73
  meter on, 94–95
  as power source, 74
  production of, 74
  signature of, 77
AC generator, 27, 74–75, 78–79
  See also Alternator; Generator.
AC power distribution, 260–64, 271
AC power generation, 258
AC voltage, 76–78
AC waveforms, 77–78, 79
Acoustics, **218**–19
Acrylic lens, **280**
Active devices, 14, **112**–24
  comparison of, 116–18
  Edison's contribution to, 112–13
  Fleming valve, 113, 114
  integrated circuits, 16–17, 119–22
  newer, 116
  oscillation, 122–24
  triode vacuum tube, 114–16
  as variable resistances, 119
Acupuncture, **159**–60
Air capacitors, 63
Aldrin, Buzz, 170
Alexanderson alternator, 13–14, 126, 130
Alexanderson, Ernst, 13–14, 126, 130
Alkaline cell, **105**
Alloy, **43**, 51
Alnico, 54
Alpha particles, 150
Alternating current, **71**
  See also AC.
Alternator, **9**, 27, **257**
  Alexanderson, 13–14
  automobile, 74–75
  commercial, 74–75
  generator versus, 74
  improving efficiency of, 78–79
  invention of, 9, 13–14
Aluminum, 22
AM (amplitude modulation), **131**, 132, **186**
AM broadcast stations, 186–88
  advantages of, 187
  classification of, 186–87
  jobs in, 187–88
AM/FM tuners, 217–18
AM stereo, 199
Amateur radio, 200–02
American Radio Relay League, 201–02
Ammeter, **56**–57, 96–97
Amperes, **25**, **26**
Amplification, **112**
Amplifier, 213–15
  differential, 153
  standards for, 215
Amplitude, **131**
Amplitude modulation. *See* AM.
Analog, **116**–17
Analog computers, **241**
Analog form of sound waves, **219**
Analog meter, **101**
Anemometer, **279**
Apollo space missions, 170
Arecibo, 109
Arithmetic logic unit (ALU), **242**
Armature, 81–82
Armstrong, Edwin, 126, 188–90
Armstrong, Neil, 170
Artificial elements, 21
Atom(s), **20**
  mass of, 21
Atomic structure, planetary theory of, 21
Atomic theory of matter, 20–21
Attractive force, 20
Audio frequencies, **59**

Audio frequency transformer, structure of, 83
Audio system components, 213–19
  accessories, 218
  AM/FM tuners, 217–18
  amplifiers, 213–15
  loudspeakers, 215–16, 217
  microphones, 217
  pickup cartridges, 216–17
  room acoustics, 218–19
  See also Sound recording and reproduction.
Audion, 115, 207
  See also Triode vacuum tube.
Automotive electronics, 303

B-scan, 152
Babbage, Charles, 239–40
Babbage's analytic engine, 239–40
Baffle, 216
Bandwidth, **226**
Bardeen, John, 14–16
Barium titanate, 29
Bar magnet, 51, 52, 53, 54
Base, **118**, 119
BASIC, 246
Battery, **103**, **104**
  charging, 106–07
  chemical reactions in, 28
  designations for, 118
  as source of electrical current, 25
  See also Cell.
Bell, Alexander Graham, 10–11, 208
Bell Laboratories research, 14–16
Berliner, Emile, 208–09
Berliner gramophone, 208–09
Beta particles, 150
Biomedical equipment technician (BMET), 161
Biophysics, **137**
Bipolar transistor, **16**, **118**
Bits, **244**
Blood pressure, measurement of, 139
Bohr, Niels, 21
Brattain, Walter, 14–16
Buffering, **246**
Busses, **246**
Bytes, **244**

Cable television, 194–99
  geosynchronous satellites in, 195–97
  installation of, 197
  pickup and distribution system for, 196
  satellite reception of, 197–99
Calculators, electronic, 249–50, 299–300
Cameras, automated, 277–79, 301–02
Capacitance, 61–66
  calculation of, 68
  factors determining, 61
  units of, 63
Capacitive reactance, 65–**66**
Capacitors, 50, **61**
  charging and discharging, 62–63
  electrolytic, 63, 64
  fixed, 63
  inductors and, 67–68
  naming, 63
  in parallel, 65
  precautions in connecting, 65
  in series, 64–65
  variable, 62
  voltage rating of, 64
Carbon, 20, 42
Carbon resistor, **43**, 44
Carbon-zinc cell, 104–05
Carey, C. R., 190
Cassettes, **219**, **226**
CAT scanner, **146**–48
  computer and, 148
  construction of, 147–48
  cost of, 148
  future of, 295–96
  tomography and, 147
Cathode-ray tube, **14**, **136**
CATV. *See* Cable television.
CB radio, 202–03
Cell, **103**–06
  alkaline, **105**
  carbon-zinc, 104–05
  miniaturized, 105–06
  primary, **105**
  rechargeable, 105
  secondary, **105**
  Volta's, 103–04
Central processing unit (CPU), **240**, 242–43
Centrifugal force, 20
Ceramic capacitors, 63
Ceramic magnets, **54**
Chemical reactions, as source of electrical pressure, 27–28
Circuit, **24**–26
  current flow in, 25–26
  open, **38**
  oscillatory, 122–24
  parallel, **40**–41
  parallel-resonant, **67**, 122–24
  reading power in, 101–02
  series, **34**, **39**–40
  series-parallel, **41**–42
  short, **39**

tuning and tuned, 67–68, 122–24
See also Integrated circuit.
Circuit breaker, **39**
Circuit diagram, integrated circuit on, 121
Citizens Radio Service, 202–03
Clocks, electronic, 250–53, 300
Cobalt, 51
COBOL, 246
Coherent light, 5, 139, **229**
Coherer, 11
Collector, 119
Color-coded resistors, 45–46
Color television, 193–94
Columbia space shuttle, 176, 177
Communications. *See* Electronic communications; Telecommunications.
Commutator, 81, 82
Compact disc (CD), **220**–21
Compass, 9, 81
Complementary symmetry, 85
Compounds, **24**
Computer(s), 239–48
analog, **241**
busses, 246
CAT scanning and, 148
central processing unit, 240, 242–43
design of, 241–42
digital, **241**–48
future of, 299, 300
historical perspective on, 239–40
home, 247–48
input-output circuits, 245–46
microprocessors and, 248–49
transistors and, 240
See also Word processing.
Computer disc, 245
Computer language, **246**
Computer memory, 243–45
Condensers, 61
See also Capacitors.
Conductivity, 21–22, 42
Conductor
electrical, **21**, 24
resistance of, 42–43
Continuity test, **98**
Control, 45
Copiers, electrostatic, 284–85, 286
Copper, 22
Coronary care equipment, 141–42
Cosmic rays, 126
Coulomb, Charles, 32–33, 53
Coulomb, **63**
Coulomb's Law of Charges, **32**–33
Counter EMF, **57**
CPU. *See* Central processing unit.
Crookes, Sir William, 136
Curie, Irene, 136
Curie, Jacques, 136
Curie, Marie, 136
Curie, Pierre, 136
Current, 21, **25**–26
Kirchhoff's Law of, 33
strength of, 95–96
See also AC; DC.
Customer service, 6
Cycle, **72**

Damping, meter, 92
D'Arsonval, Arsene, 90, 137–39, 157
D'Arsonval meter, 89, **90**–98
AC and, 94–95
connecting voltmeters and ammeters, 96–97
construction of, 90–91
development of, 137–38
measuring current strength with, 95–96
measuring electrical pressure with, 93–94
measuring heavy currents with, 96
measuring resistance and continuity with, 97–98
movement of, 91
operation of, 92–93
sensitivity of, 93
DC, 72–73
AC versus, 72–73
conversion to AC, 265–66
pulsating, **73**
DC generator, 56
DC motor, 82
Defibrillator, 140, 141
Deflection, exactly full, 93
De Forest, Lee, 2, 14, 114–15, 126, 188, 207
Demand pacemaker, 160
Democritus, 8
Detector, **113**
Diagnostic equipment, 143–57
CAT scanners, 146–48
electrocardiograph, 154–55
electroencephalograph, 152–53
electromyograph, 153–54
electron microscope, 156–57
NMR scanners, 148–49
thermographic equipment, 155–56
ultrasound, 151–52
X-ray machines, 143–46
Diathermy, **138**
development of, 138
for therapy, 157–58, 296

Dictaphone, 208
Dielectric, **61**, 62
Differential amplifiers, 153
Digital computers, **241**–48
Digital form of sound waves, **219**
Digital recording, 219–21
Digital voltmeter, **101**
Diode, **75**
Diplexer, **193**
Disc drive, **247**
Discs. *See* Floppy disc; Record discs; Video disc recording and playback.
Distortion, **77**, **214**
Dividing down, **233**
Domains, **53**
Dopants, 16
Doping, 116
Dot-matrix printer, **247**
Drain, 119
DX operators, **201**
Dynamic memory, 244
Dynamic microphone, **217**
Dynamic range, **219**
Dynamo, **257**

Echocardiography, **152**
Eddy currents, **85**
Edison, Thomas, 27, 112–13, 145, 190–91, 207–08, 257–58
Edison phonograph, 208
Effective value, 79–80
Einthoven, Willem, 154
Electric motors, 81–82
Electric shock, 138–39
Electrical charges, Coulomb's Law of, 32–33
Electrical circuit. *See* Circuit.
Electrical conductor, **21**, 24
Electrical current. *See* AC; Current; DC.
Electrical power, 26–27, 256–72
  AC, 258, 260–64, 271
  alternate sources of, 268–71
  DC, 264–66
  distribution of, 259–66, 271
  future of, 300–01
  generation of, 258–59
  historical perspective on, 257–58
  hydroelectric, **259**
  nuclear, 19, 23–24, 266–68
  solar, 268–69
  steam-produced, 262
  wind, 269–71
  See also Power.
Electrical pressure
  generation of, 27–29
  measurement of, 93–94
Electrical relay, 10
Electricity, **8**
  cost of, 263
  ions and, 22
  residential use of, 263
Electrocardiograph (EKG), **139**, **154**–55
Electroencephalograph (EEG), **139**, **152**–53
Electrolyte, **28**, **64**
Electrolytic capacitors, 63, 64
Electromagnet(s), **55**
  inventions based on, 10
  strength of, 9–10, 55–56
Electromagnetic field, **55**, 58
Electromagnetic pickups, 233–34
Electromagnetic spectrum, 126–27
Electromagnetism, 27, 55
Electromotive force. *See* EMF.
Electromyograph (EMG), **153**–54
Electron(s), **20**
  attractive force and, 20
  free, 21–22
  motion of, 25–26
Electron microscope, 156–57
Electron theory, 27
Electronic calculators, 249–50, 299–300
Electronic clocks, 250–53, 300
Electronic communications, 125–32
  conveying information by, 130–32
  historical perspective on, 125–26
  radio waves, 127–29
  sunspots and, 129–30
  See also Telecommunications.
Electronic equipment, sales of, 3
Electronic guitar, 232, 233
Electronic Industries Association (EIA), 45
Electronic organ, 232–33, 299
Electronic piano, 234–35
Electronic watches, 253, 300
Electronic weather stations, 279–80
Electronics, **8**
  automotive, 303
  in business and industry, 284–89, 302
  future of, 292–303
  in home, 275–83, 301–02
  in music, 231–35
  trends in, 289–90
Electronics Industry
  future of, 293–95
  growth of, 2–3, 293–95

history of, 2, 7–17
job opportunities in, 3–5, 7, 293–95
organization of, 3
plant organization in, 6
Electrostatic copiers, 284–85, 286
Electrostatic field, 62–63
Elements, **20**–21
artificial, 21
physical characteristics of, 20–21
Emergency room equipment, 140–41
EMF (electromotive force), 25
counter, **57**
Emitter, 119
Engineers, qualifications for, 4, 5
ENIAC, 240
Enterprise space shuttle, 177
Entrance box, **41**
Equalization, **214**
Equipment servicers, qualifications for, 4–5
Erase head, **212**
Ewing's Theory of Magnetism, 53
Exactly full deflection, 93

Farad, **63**
Faraday, Michael, 9, 27, 74, 265
Federal Communications Commission (FCC), 14
Fessenden, Reginald, 13, 126
FET voltmeter, 99–101
Fidelity, 214
Field, **192**
Field-effect transistor, **118**
Field winding, 82
Filter capacitors, 63, 64
Filtering, **50**
Fission, **23**, **267**
Fixed capacitors, 63
Fleming, Ambrose, 113
Fleming valve, 113, 114
Floppy Disc, 245
Fluoroscopy, **145**
FM (frequency modulation), **132**, **188**
FM broadcasting, 188–90
classification of stations, 190
development of, 188–90
growth of, 190
Force, lines of, **52**
Formant filters, **234**
FORTRAN, 246
Frame, **192**
Franklin, Benjamin, 8–9, 26
Free electrons, 21–22
Frequency
audio, **59**
radio. *See* Radio frequencies.
standard, 77
Frequency modulation. *See* FM.
Frequency response, and video recording, 225–26
Fresnel lens, **280**
Friction, as source of electrical energy, 28
Fuse, **39**
Fusion, **23**–24, **266**–67

Galvanometer, **90**–91
See also D'Arsonval meter.
Gamma rays, 126, 150
Gate, **100**, **118**, 119, **242**
Gemini satellites, 170
General Electric Company, 287–88
Generator, **9**, **257**
AC, 27, 74–75, 78–79
alternator versus, 74
commercial, 74–75
DC, 56
improving efficiency of, 78–79
See also Alternator.
Geosynchronous satellite, 5, **179**–80, 195–97
Germanium, research on, 15–16
Glenn, John H., 170
Goldmark, Dr. Peter, 210
Gramophone, 208–09
Gravitational force, 20
Gray range, 190
Grid-controlled rectifier, **264**
Ground, 67–**68**
Ground-fault locator, 161
Ground wave, **127**
Guitar, electronic, 232, 233

Hammond, John Hays, 232
Hard wiring, **246**
Harmonics, **77**, 234
Headphone, 110
Health occupations. *See* Medical electronics.
Heart pacer, **160**–61
Heat, as source of electrical energy, 28
Helical scan recorders, **226**–28
Helium, 20
Henry, 61
Henry, Joseph, 9–10, 55, 61
Hertz, Heinrich, 11, 13, 125
Hewlett-Packard, 4, 7
High fidelity, **214**
Hill-and-dale recording, **208**

Hillier, James, 157
Hoff, Ted, 169, 240
Home computers, 247–48
Home electronics, 275–83
  cameras, 277–79, 301–02
  color projection television, 280–81
  cordless telephones, 282–83
  future of, 301–02
  microwave ovens, 275–76
  smoke detectors, 277
  solid-state laundry units, 276–77
  video games, 281–82, 283
  weather stations, 279–80
Horseshoe magnet, 52
Hydroelectric power, **259**
Hydrogen, 20, 23

IBM, 240
IC. *See* Integrated circuit.
IC memory, 244
Iconoscope, **191**
IF (intermediate frequency), **193**
Impedance, 72, **80**–81
Impedance matching, 85
Indicating devices, 110–11
Inductance, **56**–61
  calculation of, 68
  measurement of, 61
  mutual, **59**, 60
  self-, **59**
Inductive reactance, **58**–59
Inductors, 50, **55**, **56**, 67–68
Infrared frequencies, 127
Input-output (I/O) circuits, **245**–46
Integrated circuit (IC), **119**–22
  advantages of, 119–20
  on circuit diagrams, 121
  development of, 16–17
  manufacturing of, 120, 122
  types of transistors, 121–22
  very large scale, 218
Intel, 240
Intensive care equipment, 141
Interfaces, **246**
Interlaced scan, **192**
Interlock, **275**–76
Intermediate frequency (IF), **193**
Ion(s), **22**
Ionization chamber smoke detector, **277**
Ionosphere, **127**, 130
Iron, 51

Job opportunities, 3–5, 7, 293–95

Keyboard, **247**
Kilby, Jack, 16–17
Kilobyte, **244**–45
Kilowatt-hour, **37**
Kinescope, **244**
Kirchhoff, Gustav, 33–34
Kirchhoff's Law of Current, **33**
Kirchhoff's Law of Voltage, **34**
Klystrons, 167

Laminations, **83**
Landsat D, 181–82
Laser, **139**–40, 229
Laser beam, **5**
Lateral recording, **208**
Laundry units, solid-state, 276–77
LCD (liquid crystal display), 248, **250**, 252–53
LED (light-emitting diode), 111, 112, **250**, 252–53
Life-support equipment, 140, 160–61
Light
  coherent, 5, 139, **229**
  as source of electrical energy, 28
Light-emitting diode (LED), 111, 112, **250**, 252–53
Light splitter, **230**
Lines of force, **52**
Liquid crystal display (LCD), 248, **250**, 252–53
Lithium, 20
Load, 24
Lodestones, 8
Loop law, 34
Loudspeaker, 85, 86, 110, **215**–16, 217
Luminescence, **150**

Magnet(s)
  Alnico, 54
  bar, 51, 52, 53, 54
  ceramic, **54**
  as complete systems, 53
  formation of, 54
  horseshoe, 52
  natural, 8, 51
  permanent, **51**–52, 54
  temporary, **51**
  weakening of, 54
  See also Electromagnet.
Magnetic core memory, 244
Magnetic disk recording, 245
Magnetic field, **52**
  behavior of, 52–53
  of Earth, 50, 51
  See also Electromagnetic field.
Magnetic materials, 51

Magnetic patterning, 245
Magnetic poles, 50–51, 53
Magnetic recording and playback
audio, 211–12, 219
video, 224–28
Magnetism, 50, 53
See also Electromagnetism.
Magnetron vacuum tube, **275**
Mammography, **156**
Manganin, 42
Marconi, Guglielmo, 11–12, 14, 125, 127
Mark I computer, 240
Matter, atomic theory of, 20–21
Maxwell, James Clark, 127
Medical electronics, 135–62
acupuncture, 159–60
CAT scanners, 146–48, 295–96
coronary care equipment, 141
for diagnosis, 143–57, 295–96
diathermy, 138, 157–58, 296
electrocardiograph, 154–55
electroencephalograph, 152–53
electromyograph, 153–54
electronic life-support equipment, 140, 160–61
electron microscope, 156–57
emergency room equipment, 140–41
equipment maintenance, 161
future of, 295–96
general patient care equipment, 142
historical perspective on, 136–39
intensive care equipment, 141
NMR scanners, 148–49, 295–96
operating room equipment, 139–40
overview of, 139–42
for therapy, 157–60, 296
thermographic equipment, 155–56
ultrasound equipment, 151–52, 158–59, 295, 296
X-ray machines, 143–46
Megawatt, **37**
Megohms, **46**
Memory, computer, 243–45
Mercury arc valve, **264**–65, 266
Metafine, 219
Meters. *See* D'Arsonval meter; Volt-ohm-milliammeters.
Metglas, 271
Mica capacitors, 63
Microcomputer, 240, **241**
Microfarad, **63**
Micromassage, **159**, 296
Microphone, 110, 217
Microprocessor, 248–49, **276**, 301
Microscope, electron, 156–57
Microvolt, 152
Microwave oven, 275–76
Milliamperes, 36
Miniaturized cell, 105–06
Minicomputer, **241**
Modem, **247**–48
Modulation, **130–31**
Modules, **242**
Monochrome, **192**
Monophonic sound reproduction, **213**
Moog synthesizer, 235
Morse, Samuel, 10
Morse code, 10
Mosaic, **191**, 192
Motion pictures, sound, 209
Motors
DC, 82
electric, 81–82
shunt field, 82
universal, **81**
Multihop transmission, **129**
Multiplier, **94**
Multiplier resistor, **94**
Multispectral scanner, 182
Music electronics, 231–35
electromagnetic pickups, 233–34
future of, 298–99
organs, 232–33
pianos, 234–35
synthesizers, 235
Mutual inductance, **59**, 60

NASA electronics, 166–70
NASA space program, 169–78
early, 170
effect on electronics, 169–70
space probes, 20, 170–77
space shuttle orbiters, 177–78
National Oceanic and Atmospheric Administration. *See* NOAA.
Neptunium, 21
NESDIS (National Environmental Satellite Data and Information Service), 179
Neutral particles, 20
Neutrons, **20**
Nichrome, 42
Nickel, 51
Nipkow, Paul, 191
NMR (nuclear magnetic resonance) scanner, **148**–49, 295–96
NOAA (National Oceanic and Atmospheric Administration), 166, 178–82
Noise reduction circuits, 214–15
Noyce, Robert, 16–17
N-type silicon, 116
Nuclear fission, **23**, **267**

Nuclear fusion, **23**–24, **266**–67
Nuclear power plant, 19, 267–68

Oersted, Hans Christian, 9, 55
Ohm, **9**, **26**
Ohm, George, 9, 35
Ohmmeter, **97**–98
Ohm's Law, 9, **35**–36
Open circuit, **38**
Operating room equipment, 139–40
Optical fibers, 5
Optoelectronics, 5, **288**–89
Organ, electronic, 232–33, 299
OSCAR space satellites, **201**
Oscillation, **112**, **122**–24
Oscillator, 122
Oscillatory circuit, 122–24
Oscilloscope, **77**, **103**, 141

Pacemaker, **160**–61
Paper capacitors, 63
Parabolic reflector, **167**
Parallel circuit, **40**–41
Parallel-resonant circuit, **67**, 122–24
PDC (pulsating DC), **73**
Peak value, 79–80
Peripheral, **247**
Permanent magnet, **51**–52, 54
Permanent-magnet dynamic speaker, 216
Personal computer, 241, 247–48
Phase, 78
Phonograph, 208–09, 210
  See also Audio system components.
Phonograph pickup, **210**
Photoelectric cell smoke detector, **277**
Photographic equipment, 277–79, 301–02
Photons, 28
Photo-resist, 120
Photovoltaic cell, 268
Piano, electronic, 234–35
Pickup cartridges, 111, **216**–17
Picofarad, **63**
Piezoelectric activity, 28–29
Piezoelectric devices, 111
Pioneer Venus I and II, 171–72
Pioneer X, 176–77
Plant organization, 6
Plutonium, 21
Polar orbiting satellite, **179**, 180
Pot, 45
Potentiometer, **45**
Poulsen, Valdemar, 211
Power, 26–27, **37**
  See also Electrical power.
Power transformer, 83
Pressure. *See* Electrical pressure.
Primary, **82**
Primary cell, **105**
Production line, 6
Production workers, qualifications for, 4
Protons, **20**
Prototype, **41**
P-type silicon, 116
Pulsating DC (PDC), **73**
Pulse-echo diagnostic procedure, **151**
Pupin, Michael, 189
Push-pull circuit, **213**

Quadrature machines, **225**
Quartz, 29, 111

Radio, 127–32, 185–90
  amateur, 200–02
  AM stations, 186–88
  Citizens Radio Service, 202–03
  early, 2, 11–12, 35, 185–86
  FM stations, 188–90
  future of, 199–200, 297
  invention of, 11–12
  phonograph and, 210–11
  quartz crystal and, 29
Radio emergency associated citizens teams (REACT), 203
Radio frequencies, **59**
  AM, **131**, 132
  reading power at, 102–03
Radio waves
  direction of, 128
  in electromagnetic spectrum, 127
  propagation of, 127–29
Radioactive scanning, 150–51
Random access memory (RAM), **244**
Read only memory (ROM), **244**
Receiver trimming capacitor, 62
Receiver tuning capacitor, 62, 63
Rechargeable cell, 105
Record discs, 208–09, 210–11
Recording. *See* Sound recording and reproduction; Video disc recording and playback; Video tape recording and playback.
Recording head, **212**
Rectification, **264**–65
Relay, electrical, 10
Research and development (R&D), 6
Resistance, **9**, **22**, **42**
  calculation of, 36
  measurement of, 97–98

Resistive materials, 42
Resistors, **42**–46
  carbon, **43**, 44
  color coding, 45–46
  variable, **43**–45
  wire-wound, **43**, 44
Resolution, **226**
Retentivity, **51**
Robotics, **287**
Robots, 287–88
Roentgen, Wilhelm, 136–37
ROM (read only memory), **244**
Root-mean-square (RMS) value, 80
Rotor plates, 62
Rutherford, Ernest, 150

Salaries, 7
Salt, 24
Satellite(s), 179–82
  geosynchronous, 5, **179**–80, 195–97
  OSCAR, 201
  polar, **179**, 180
Satellite communications, 5–6
Satellite television, 197–99
Saturn V launch rockets, 170
Sawtooth wave, **234**
Scanning, **147**–51
Schematic, **46**, **93**
Scintillation detector, **150**
Secondary, **82**
Secondary cell, **105**
Selectivity, **68**
Self-inductance, **59**
Semiconductors, **15**–16
Series circuit, **34**, **39**–40
Series-parallel circuit, **41**–42
Shockley, William, 14–16
Short circuit, **39**
Shunt, **95**
Shuntfield motor, 82
Sideband, **132**
Signal, 14
Silence, zone of, 129
Silicon, 28, 116
Silicon chip, 121
Silver, 21
Sine wave, **77**, 123
Single-sideband transmission, **201**
Skip distance, **129**
Skylab, **175**–76
Sky wave, **127**, 128
Slant-track recorders, **226**
Slip-rings, 79
Smoke detector, 277
Software, **248**
Solar cell, **268**
Solar energy, 28, 268–69
Solenoids, **59**–61
Solid-state, **75**
Solid-state laundry units, 276–77
Sound-motion pictures, 209
Sound recording and reproduction, 207–21
  acoustics and, 218–19
  digital technology and, 219–21
  distortion in, **214**
  electronics and, 209–10
  equalization, **214**
  future of, 297–98
  high fidelity, **214**
  hill-and-dale, **208**
  history of, 207–11
  lateral, **208**
  magnetic tape, 211–12, 219
  monophonic, **213**
  noise reduction circuits in, 214–15
  stereophonic, **213**–14
  See also Audio system components.
Sound waves, **216**
  analog form of, **219**
  digital form of, **219**
Space exploration, 165–82
  future of, 296
  historical perspective on, 166
  NASA electronics in, 166–70
  NASA program, 170–78
  NESDIS, 179–82
  NOAA, 178–79
Space probes, 20, **170**–77
  Pioneer Venus I and II, 171–72
  Pioneer X, 176–77
  Skylab, 175–76
  Viking, 171
  Voyager I and II, 172–75
Space shuttle orbiter, **177**–78
Spacelab, **178**
Spectrum, electromagnetic, 126–27
Square wave, **159**
Standard frequencies, 77
Stator plates, 62
Steam-produced power, 262
Step-down ratio, 82–83
Step-up ratio, 82–83
Stereo sound reproduction, **213**–14, 297, 298
Stirrer, **275**
Strobe flash tube, **278**
Stylus, **217**
Sunspots, 129–30

Supervisors, qualifications for, 4
Synthesizer, music, **235**

Tape cassettes, **219**, **226**
Tape recorders
audio, 211–12, 219
video, 224–28
Taut band armature support, 91, 92
Technicians, qualifications for, 4–5
Tektronix, 7
Telecommunications, **185**–204
amateur radio, 200–02
AM broadcast stations, 186–88
cable television, 194–99
Citizens Radio Service, 202–03
FM broadcast stations, 188–90
future of, 199–200, 296–97
historical perspective on, 185–86
television broadcasting, 190–99
Telegraph, invention of, 10
Telemetering, 142, **261**
Telephone
cordless, 282–83
invention of, 11
Telephone answering system, 284
Television, 190–200
AM and FM in, 132, 193
cable, 194–99
color, 193–94
early technology, 190–91
future of, 199–200, 296, 297
invention of, 14
jobs in, 194
projection, 280–81
resistors and, 44
standards for, 192–93
system, 191–92
transformers and, 86
Temporary magnet, **51**
Terminal equipment, **245**, 246
Tesla, Nikola, 12–13, 27
Texas Instruments, 303
Therapeutic equipment, 157–60
acupuncture, 159–60
diathermy, 157–58
ultrasound, 158–59
Therapy, **157**
Theremin, Leon, 232
Theremin (musical instrument), 232
Thermionic device, 14
Thermionic emission, **116**
Thermistor, **142**
Thermocouple, 28
Thermography, **155**–56, 295
Time base, **251**
Timekeeping, electronic, 250–53
Tomography, **147**
Tracing pen, 110–11
Transducers, **110**–12
common, 110–11
efficiency of, 111–12
for measuring blood pressure, 139
Transformers, **82**–86
applications of, 85–86
audio frequency, 83
early, 12–13
efficiency of, 84–85
power, 83
power supply of, 84
step-up and step-down ratios, 82–83
Transistors
bipolar, **16**, **118**
computers and, 240
development of, 15–16
field-effect, **118**
IC, 121–22
in push-pull circuit, 213
Transmission line, **259**–61
Transmitting capacitor, 62
Transponder, **195**–97
Transverse recording, **225**
Traveling-wave tubes, 167
Triode vacuum tube, 2, 14, **114**–16, 207
Trojan Nuclear Plant, 19, 268
Tuned circuit, 67–68, 122–24
Tuners, AM/FM, 217–18
Tuning, **50**
Tuning circuit, 67–68
Turns ratio, 82
Tweeters, **216**

Ultrasonic frequencies, 127
Ultrasonic waves, **59**
Ultrasound, **151**
for diagnosis, 151–52, 295
for therapy, 158–59, 296
Ultraviolet frequencies, 126
Universal motor, **81**
Uranium, 20, 21, 267

Vacuum capacitors, 63
Vacuum tubes
magnetron, **275**
sound recording and, 209–10
for space-vehicle transmitters, 167
triode, 2, 14, **114**–16, 207
as a variable resistance, 115–16
Variable capacitors, 62

Variable resistors, **43**–45
Varistor, **278**
Very large scale integrated circuits (VLSICs), 218
Video disc recording and playback, 229–31, 298
Video games, 281–82
Video tape recording and playback, 224–28
  Beta system, 227
  compatibility and, 226–27
  equipment sales, 228
  frequency response and, 225–26
  future of, 298
  helical scan recorders, **226**–28
  for home use, 227–28
  quadrature machines, 225
  VHS system, 227
Viking mission, 171
Voice synthesizing, 303
Volt, **27**
Volta, Alessandro, 9, 27, 103–04
Voltage, 25
  AC, 76–78
  calculation of, 35–36
  Kirchoff's Law of, 34
Voltage drop, **34**
Voltage rating, **64**
Voltmeter
  analog, **101**
  connecting with ammeter, 96–97
  digital, **101**
  FET, 99–101
Volt-ohm-milliammeter (VOM), **98**–103
  digital, 101
  FET, 99–101
  oscilloscope, 103
  reading dial scale, 98–99
  reading power at radio frequencies, 102–03
  reading power in circuit, 101–02
Volume unit (VU) meter, **212**
VOM. *See* Volt-ohm-milliammeter.
Voyager I and II, 172–75

Wages, 7
Watches, electronic, 253, 300
Watson, Thomas, 11
Watt, **27**, **37**
Watt, James Stephenson, 37
Wattage rating, **43**
Watt-hour, **37**
Wattmeter, 37, 101, **102**
Waveguide, **275**
Weather satellites, 179–81
Weather stations, electronic, 279–80
Westinghouse, George, 27, 82, 258
Wind power, 269–71
Wireless, **185**
Wire-wound resistor, **43**, 44
Woofers, **216**
Word processing, 285–86, 287
Writing speed, **224**

Xerography, **285**
Xerox Corporation, 285
X-rays, 126, **137**
  control equipment for, 144–45
  dental, 145–46
  for diagnosis, 143–46
  early use of, 137
  fluoroscopy and, 145

Zone of silence, 129
Zworykin, Vladimir, 14, 157, 191

Document generated by Anna’s Archive around 2023-2024 as part of the DuXiu collection
(https://annas-blog.org/duxiu-exclusive.html).

Images have been losslessly embedded. Information about the original file can be found in PDF attachments. Some stats (more in the PDF attachments):
{
  "filename": "NDAzODUwMjMuemlw",
  "filename_decoded": "40385023.zip",
  "filesize": 98003109,
  "md5": "b355125d38afc6215aebb89afd229458",
  "header_md5": "bd2a0c46b78e52f76b687cc70adc2713",
  "sha1": "ccf537b6a8c9dc9cdf262313a91efc255560521b",
  "sha256": "bca4b80774be959f2e2956a5f54f2fb9df4645f4e1e186d556994b041e5fd30c",
  "crc32": 1827991917,
  "zip_password": "52gv",
  "uncompressed_size": 109270735,
  "pdg_dir_name": "\u00ed\u2562ELECTRONICS CONCEPTS\u00fa\u00bcAPPLICATIONS\u00fa\u00bcAND HISTORY\u00ed\u2556_40385023",
  "pdg_main_pages_found": 315,
  "pdg_main_pages_max": 315,
  "total_pages": 329,
  "total_pixels": 2002320320,
  "pdf_generation_missing_pages": false
}